TU WIEN Informatics

Untersuchen der Auswirkung von Tumorsegmentierungen auf die Radiomics-Analyse durch Visual Analytics

DIPLOMARBEIT

zur Erlangung des akademischen Grades

Diplom-Ingenieurin

im Rahmen des Studiums

Medizinische Informatik

eingereicht von

Michelle Duong, BSc
Matrikelnummer 01406194

an der Fakultät für Informatik

der Technischen Universität Wien

Betreuung: Assistant Prof. Dr. Renata Georgia Raidou

Wien, 15. Jänner 2023

Michelle Duong	Renata Georgia Raidou

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

TU WIEN Informatics

Investigating the Effect of Tumor Segmentations on Radiomics Analysis through Visual Analytics

DIPLOMA THESIS

submitted in partial fulfillment of the requirements for the degree of

Diplom-Ingenieurin

in

Medical Informatics

by

Michelle Duong, BSc
Registration Number 01406194

to the Faculty of Informatics

at the TU Wien

Advisor: Assistant Prof. Dr. Renata Georgia Raidou

Vienna, 15th January, 2023

Michelle Duong	Renata Georgia Raidou

Technische Universität Wien
A-1040 Wien ▪ Karlsplatz 13 ▪ Tel. +43-1-58801-0 ▪ www.tuwien.at

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Erklärung zur Verfassung der Arbeit

Michelle Duong, BSc

Hiermit erkläre ich, dass ich diese Arbeit selbständig verfasst habe, dass ich die verwendeten Quellen und Hilfsmittel vollständig angegeben habe und dass ich die Stellen der Arbeit – einschließlich Tabellen, Karten und Abbildungen –, die anderen Werken oder dem Internet im Wortlaut oder dem Sinn nach entnommen sind, auf jeden Fall unter Angabe der Quelle als Entlehnung kenntlich gemacht habe.

Wien, 15. Jänner 2023

Michelle Duong

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Danksagung

Zuallererst möchte ich meiner Betreuerin Renata Raidou für ihre unermüdliche Unterstützung und ihr Engagement meinen tiefsten Dank aussprechen. Ihre brillanten Vorschläge und ihr Fachwissen haben mir nicht nur in jeder Phase des Schreibens dieser Arbeit, sondern auch während meines gesamten Masterstudiums geholfen.

Außerdem möchte ich mich bei meinen Eltern, meiner Schwester Jessica und meinem Partner Tamas bedanken. Ich weiß ihre emotionale Unterstützung, ihre ermutigenden Worte und ihr Verständnis während dieser Zeit zutiefst zu schätzen. Schließlich möchte ich auch all meinen Freunden, die mich während meines gesamten Studiums unterstützt haben, danken.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acknowledgements

First and foremost, I would like to express my deepest appreciation to my supervisor Renata Raidou for her unwavering support and commitment. Her brilliant suggestions and expertise tremendously guided me not only at every stage of writing this thesis but also throughout my master's program.

Further, I wish to extend my gratitude to my parents, my sister Jessica, and my partner Tamas. I deeply appreciate their emotional support, words of encouragement and understanding during this journey. Lastly, I would also like to thank all my friends who supported me through all my studies.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Kurzfassung

In den letzten Jahren hat Radiomics die klinische Beurteilung von Tumoren revolutioniert. Durch das Extrahieren von quantitativen Merkmalen aus medizinischen Bildern bietet dieser Ansatz eine objektive Analyse von Tumorgewebe, was letztlich medizinischen Expertinnen und Experten bei Entscheidungsfindungsprozessen in Bezug auf Diagnose und Behandlung hilft. Radiomics ist jedoch in hohem Maße von der Qualität der Tumorsegmentierung abhängig. Unterschiedliche Tumorabgrenzungen, die sich aus der Variabilität innerhalb und zwischen den Beobachterinnen und Beobachtern ergeben, können die Ergebnisse der Radiomics–Analyse erheblich beeinträchtigen. Unseres Wissens wurde bisher noch nicht untersucht, wie sich die Unterschiede zwischen den Beobachterinnen und Beobachtern bei der Tumorsegmentierung auf die Radiomics-Analytik auswirken.

Diese Arbeit untersucht, wie verschiedene Tumorsegmentierungen die Radiomics–Analyse beeinflussen. Wir entwickeln daher das Visual Analytics–Tool *ProSeRa (**Pro**babilistic **Se**gmentation on **Ra**diomics, übersetzt: Probabilistische Segmentierung auf Radiomics*), das Visual Analytics zur Erforschung der Auswirkungen probabilistischer Tumorsegmentierung auf Radiomics bietet. Wir befähigen die Benutzerinnen und Benutzer dazu, die Ergebnisse unserer Radiomics-Analyse mit Bezug auf klinischen Daten in Anlehnung an Schwellenwerten für die Segmentierungsgenauigkeit, die wir anhand der Übereinstimmung der Beobachterinnen und Beobachter berechnen, zu untersuchen. Wir bieten Möglichkeiten zur Erkundung und Analyse der Radiomics-Daten, u.a. mithilfe von Algorithmen zur Dimensionalitätsreduktion und Mechanismen zur Clusteranalyse in Verknüpfung mit effektiven und aussagekräftigen Visualisierungen. *ProSeRa* erleichtert die Bewertung der Robustheit der Radiomics-Analyse und unterstützt die Erforschung der Auswirkungen der Segmentierung auf die Analyse.

Basierend auf der Evaluierung unserer Ergebnisse können wir schlussfolgern, dass, wie erwartet, die Variabilität bei Tumorsegmentierung die Ergebnisse der Radiomics-Analyse erheblich beeinflusst. Die Auswirkung war besonders deutlich in der Clusteranalyse, die unterschiedliche Ergebnisse für verschiedene Schwellenwerte der Segmentierungsgenauigkeit geliefert hatte. Dabei haben wir festgestellt, dass zusätzliche Variablen, wie z.B. das Gesamttumorstadium, für die Gruppierung von Patienten in Cluster entscheidend sind.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Abstract

In recent years, radiomics has revolutionized the clinical assessment of tumors. By extracting quantitative features from medical images, this approach provides an objective analysis of tumorous tissues, which ultimately aids medical experts in decision-making processes regarding diagnosis and treatment. However, radiomics is highly dependent on the quality of tumor segmentation. Different tumor delineations resulting from intra- and interobserver variability may significantly affect the results of radiomics analysis. To our knowledge, no prior research has been conducted on the impact of interobserver differences in tumor segmentations on radiomic analytics.

This thesis aims to investigate how different tumor segmentations influence radiomics analysis. We therefore design and propose the visual analytics tool *ProSeRa (**Pro**babilistic **Se**gmentation on **Ra**diomics*), which provides visual analytics strategies for exploring the impact of probabilistic tumor segmentation on radiomics. We empower the users to examine the results of our radiomics analysis with respect to clinical data based on segmentation accuracy thresholds, which we calculate based on the observers' agreement. We provide ways to explore and analyze the radiomics data using, among others, dimensionality reduction algorithms and cluster analysis mechanisms in conjunction with effective and expressive visualizations. *ProSeRa* facilitates the assessment of the robustness of the radiomics analysis and supports the exploration of the impact of segmentation on the analysis.

Based on the evaluation of our results, we conclude that, as anticipated, variability in tumor segmentations considerably influences the radiomics analysis results. The impact was especially prominent in the cluster analysis, which provided different outcomes for different segmentation accuracy thresholds. Thereby, we detected additional variables, such as the overall tumor stage, being crucial for grouping patients into clusters.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Contents

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 1

Introduction

Radiomics is an emerging field which deals with extracting a large number of quantitative features from medical images [MML+20]. By means of these so-called *radiomic features*, the characteristics of tumors such as shape and tissue heterogeneity are assessed, with the scope of supporting the decision-making process and the prediction of treatment outcomes. This approach thus complements the conventional workflow of radiologists and oncologists qualitatively describing tumor phenotypes, since it additionally provides computational algorithms to objectively analyze tumors. Radiomics has been proven to boost the task of uncovering patterns which may have been previously unknown to the naked eye [MML+20, YA16].

A challenging problem which arises in this domain is the high dependency of radiomics extraction performance on the quality of image segmentation. In oncology specifically, segmentation requires the delineation of tumors, which can be manually conducted by radiologists or (semi-)automatically with sophisticated algorithms. Since automatic segmentations are not completely robust yet, manual segmentations are considered as the state of the art [RBR+18, TMB+18, vTCTL+20]. As this is usually done slice-by-slice, however, it is incredibly cumbersome and time consuming. Another problem of manual delineations is reproducibility: experts also may perceive tumor borders differently on medical images, resulting in both intra- and interobserver differences. This again may lead to a considerable variability in the extracted radiomics feature values, the computation of which relies on the segmentation outcome [SvCT+20].

To our knowledge, no research to date has investigated the aspect of interobserver variability and, more generally, tumor segmentation accuracy in the context of radiomics. Previous research did both investigate segmentation variability and propose approaches for radiomics analysis, but focused on these topics separately. It is therefore in our strong interest to explore the effect of tumor probability on the radiomics analysis, as it may provide valuable information on the robustness of radiomics as part of the workflow of

tumor analytics. We aim to execute this through visual analytics to provide deep insight onto radiomics data and their dependency on segmentation accuracy.

1.1 Aim of this Work

In this master thesis, we examine the following research question:

> *"To what extent does variability in tumor segmentations and the consideration of segmentation probability influence the outcome of radiomics analysis?"*

We answer this by developing a visual analytics tool which supports the exploration of radiomics data with respect to clinical data and probabilistic tumor segmentation. We particularly investigate the effect of interobserver variability in tumor segmentations by resorting to a data set containing tumor delineations by multiple radiation oncologists [WAKD19]. Our goal is to empower the user to flexibly analyze the impact of different segmentations on the radiomics analysis – also within a probabilistic context. Based on a varying set of tumor segmentations, we further analyze which radiomic features are highly relevant for predicting a good segmentation as close as possible to available ground truth. Our visual analytics tool is expected to support the users to fulfill the following three tasks:

A Visualize and analyze radiomic features together with respective clinical data (ground truth)

B Investigate the effect of probabilistic tumor segmentations on the radiomic features analysis

C Provide user interaction to assess the radiomics analysis of probabilistic tumor segmentations with respect to clinical data

1.2 Contributions

In order to fulfill the aims of our thesis, we provide the interactive visual analytics tool *ProSeRa (**Pro**babilistic **Se**gmentation on **Ra**diomics)*. This tool is a web application which aims to facilitate examining the effects of tumor segmentations on the radiomics analysis. By means of interactive visualizations, it intends to offer insight into the correlation between segmentation accuracy with radiomics and clinical data.

The main contributions of this thesis can be summarized as follows:

- We provide an interactive visual analytics tool which supports the exploration of the **impact of probabilistic tumor segmentation on radiomics analytics**.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

- We design **visualizations representing radiomics information together with clinical data** to encourage the identification of correlations and interesting patterns within a probabilistic tumor segmentation context.

We base our findings on the results of the radiomics analysis for the "NSCLC-Radiomics-Interobserver1" dataset provided by The Cancer Imaging Archive (TCIA)[1]. We have found this dataset to be suitable for our research as it contains tumor segmentations by multiple radiation oncologists as well as the respective medical images and clinical data. We performed radiomics analysis on the data by extracting radiomic features before augmenting our information space with outcomes from automated analysis methods. Following the visualization of the radiomics analysis data, we integrated the aspect of probabilistic tumor segmentation and investigated the impact of the latter on the former. This involved computing segmentations based on the interobserver variability in our input data followed by the generation of radiomics data for these.

We evaluate our results using Visual Data Analysis and Reasoning (VDAR) [LBI+12] and Qualitative Result Inspection (QRI) [IIC+13] as our evaluation methods. Whereas the first determines the quality of data analysis, the latter focuses on the qualitative discussion of the findings. We highlight the unexpected but also anticipated outcomes by means of usage scenarios. Our results generally show a considerable influence of tumor segmentation probability on the radiomics analysis. Based on these, we also state the limitations of our research.

1.3 Structure

Our thesis will be structured as follows:

Chapter 2 conveys background knowledge about interobserver variability and the field of radiomics in order to ensure a better understanding of related topics discussed in the following chapters.

Chapter 3 presents the state-of-the-art of known algorithms and methods for probabilistic tumor segmentation and available visual analytics tools for radiomics data.

Chapter 4 explains the methodology applied concerning the usage of materials, design of user interfaces, and the implementation of the visual analytics tool.

Chapter 5 involves the revelation and evaluation of the results followed by the discussion of the main findings and statement on the limitations of this thesis.

Chapter 6 concludes the thesis by summarizing the main findings of the thesis and stating the prospects of future challenges.

[1]TCIA: `https://www.cancerimagingarchive.net`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 2

Theoretical Background

Tumor assessment using radiomics is a practice which seeks to improve the prognosis and diagnosis by means of sophisticated algorithms. While this process is considered objective, the segmentation of tumors required for the radiomics extraction is regarded subjective due to this task commonly being done manually. Hence, the results of the radiomics analysis may be heavily impacted by segmentation variability caused by different human perceptions of tumors on medical images. This chapter discusses the issue of interobserver variability in tumor tissue identification and explains the radiomics method including its workflow.

2.1 Interobserver Variability

In the medical context, interobserver variability refers to the disagreement of tissue identification among observers. This especially occurs in the manual segmentation of tumors because of different perceptions of tumor borders on medical images (Figure 2.1). Reasons for this variation comprise the misinterpretation of similar conditions or the difference in identification criteria for the condition. These findings were concluded by Watadani et al. [WSJ+13] in their study of quantifying interobserver variability among radiologists outlining honeycombing in the lungs. Their results show solely moderate agreement among readers, regardless of coming from the same geographic region, having the same experience level or specializing in the same subspeciality. Saha et al. [SGH+16] also came to a similar conclusion in their study of analyzing interobserver variability among readers assessing breast tumor tissues. They also observed a moderate variability and measured 0.6 of average agreement regarding overlap using the Dice coefficient. This value could be raised to 0.77 when incorporating automatic segmentation algorithms in addition to the manual process.

Proper tumor segmentations are crucial for an accurate tumor diagnosis. The information obtained from the delineations are in fact significant markers for treatment planning and

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

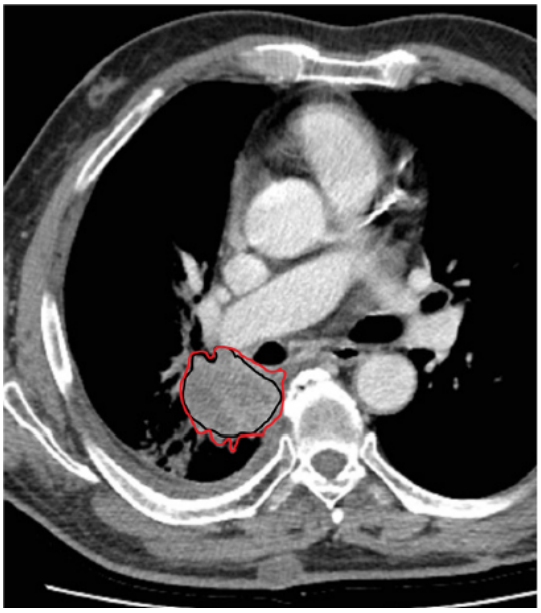


Figure 2.1: Interobserver variability present in the manual segmentation of a lung tumor, with red and black lines representing the tumor delineations conducted by one expert each [RBR+18]

prediction outcome for cancer patients. As a result, interobserver variability in tumor delineations would considerably influence the assessment of tumors. This again would impact the decision-making process and therapy outcome [DS16, EK80].

Tumor delineation also constitutes an important aspect in the effect of stability of radiomic features. Radiomics metrics may be affected by interobserver variability, resulting in considerable variation in accuracy [HLD+17]. In their study of measuring the effects of interobserver segmentation variabilities, Wong et al. [WBW+21] concluded notable differences in the prediction of certain mutations related to pancreatic cancer based on various tumor segmentations.

2.2 Radiomics

Radiomics is an emerging field of research which regards the extraction of quantitative features from clinical images. It aims to support the clinical decision-making process by finding correlations between both medical images and the respective clinical data. By means of mathematical algorithms, radiomic features such as shape are obtained, enabling the disclosure of characteristics which may not be visible to the naked eye. Due to the advantages of in-depth analyses of tumors, radiomics has been widely used for applications in personalized medicine and oncology in general such as the differentiation between benign and malign tumor cells and the prediction of treatment response [LRVL+12, MML+20, WD17, YA16].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

2.2.1 The Radiomics Workflow

The radiomics approach of extracting mineable data from clinical images involves a defined, complex process. It can be divided into the following four steps: (1) image acquisition and reconstruction, (2) image segmentation, (3) feature extraction and quantification, and (4) analysis model building. An overview of the radiomics workflow is pictured in Figure 2.2. Each stage requires a thorough evaluation in order to provide reliable and stable models for predicting outcomes [RBR+18, TMB+18].

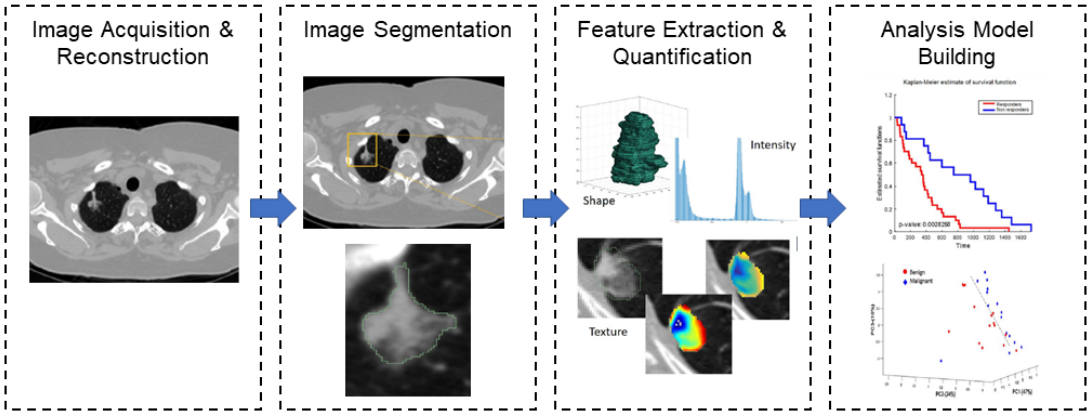


Figure 2.2: Overview of the radiomics workflow consisting of four steps: image acquisition and reconstruction, image segmentation, feature extraction and quantification, and analysis model building (based on Rizzo et al. [RBR+18] and Thawani et al. [TMB+18])

Step 1: Image Acquisition and Reconstruction

Image acquisition is part of a typical clinical routine for diagnosing medical conditions. Techniques represent a variety of parameters, including spatial resolution and number of excitations. Based on what information is required to be gathered, different imaging modalities are chosen. The most common ones include Computed Tomography (CT), Positron Emission Tomography (PET) and Magnetic Resonance Imaging (MRI). As the acquired images serve as the basis for extracting radiomics features, the value of these features are heavily dependent on variables such as image noise. Variations in reconstruction algorithms and parameters among vendors also affect the quality of feature extraction. Therefore, the probability of unstable features as a result of certain reconstruction settings should be taken into account [RBR+18, TMB+18].

Step 2: Image Segmentation

The segmentation of images is considered a crucial step of the radiomics workflow since it greatly affects the outcome of feature extraction. Generally, the delineation of the so-called Region of Interest (ROI) for two-dimensional (2D) images and Volume of Interest (VOI) for three-dimensional (3D) images is considered particularly challenging due to

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

tumors often exhibiting unclear borders. In order to overcome this difficult task, there are three ways to perform tumor segmentation: manually, semi-automatically or fully automatically [RBR+18, vTCTL+20].

Manual segmentations have the advantage of ensuring accuracy, with experts choosing the ROI/VOI by hand. However, major limitations include not just the high demand in time and efforts but also the intra- and inter-observer variability of tumor assessment. The latter also applies to segmentations done semi-automatically, which are based on standard delination algorithms such as thresholding [SV16] or region-growing [PT01]. Even though these methods allow the process to be more time-efficient, some require the expert to manually select seed points regarding their view of ROIs, resulting in bias among observers. On the contrary, fully automatic segmentation is considered relatively fast and neutral as it bypasses the intra- and inter-observer bias. No human interaction is required as it is based on deep-learning algorithms, hence relying on generalizability. However, this factor may present significant drawbacks when dealing with certain datasets, with the possibility of retrieving completely inaccurate results. Therefore, choosing the suitable segmentation method is critical for reaching the required robustness and reproducibility of features [TMB+18, RBR+18, vTCTL+20].

Step 3: Feature Extraction and Qualification

Following image segmentation, the main part of the radiomics workflow, the feature extraction, can be executed. This step deals with the calculation of radiomic features [TMB+18]. Considering that this may be done in multiple ways, it is advised to comply to the Image Biomarker Standardization Initiative (IBSI) guidelines [ZVA+20], which present a consensus on standardized calculations for radiomic features. The extraction step can be accompanied with the usage of image filters which can help improve the quality of images for a more optimized extraction. For example, the Gaussian filter may be used to reduce noise in images [RG21]. There are several open-source libraries which support the extraction of radiomics features in clinical images. One example is PyRadiomics [vGFP+17a], which is an open-source platform implemented in Python aiming for feature extraction done in an easy and reproducible way. Further information about the types of extractable radiomics features are described in Subsection 2.2.2.

Step 4: Analysis and Model Building

The fourth and final step of the radiomics process consists of selecting appropriate features for analysis purposes and ultimately building a prognostic model. This stage can generally be divided into two substeps: (1) feature selection and dimension reduction and (2) analysis and model building. The first substep concerns including reproducible features and excluding non-reproducible and non-relevant ones in order to obtain results which are generalizable. For example, features which are prone to manifest intra- and interobserver variability should not be considered into the selection. Usually, unsupervised approaches are utilized for this substep, with the most commonly known including:

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

- **Principal Component Analysis (PCA)**: A smaller set of uncorrelated variables is extracted from a larger set with correlated variables in order to show variations within the data [Jol02].
- **Cluster Analysis**: Similar features are grouped together showing redundancy and correlations of each group. Heatmaps (Figure 2.3) are often used to visualize relationships [WF09].
- **Multi-dimensional Scaling (MDS)**: Similarity or dissimilarity data are represented as distances in geometric spaces. The aim is a low-dimensional data representation which is mindful of the distances in the initial space [BG97, Kru64a, Kru64b].
- **t-distributed Stochastic Neighbor Embedding (t-SNE)**: Similar data points are transformed to joint probabilities. In doing so, it is strived for the minimal Kullback-Leibler (KL) divergence [Kul68] between these probabilities. The KL divergence, which is also known as relative entropy, is a unit of measure used to quantify the similarity between two density distributions [vdMH08].

Subsequently, supervised approaches are followed for the analysis and model building. The selected set of relevant, uncorrelated features are utilized for training the predictive model. The following algorithms are one of the most commonly used for this purpose:

- **Random Forest**: Tree predictors consisting of trees which are dependent on random vector values are combined [Bre01].
- **Linear Regression Analysis**: The relationship between the independent variable X and the dependent variable Y is shown in a coordinate system. Whereas the values for the first are represented by dots, the latter is illustrated as a linear line known as the regression line [SHB10].
- **Neural Networks**: Imitating the neurons of a brain, neural networks consist of model neurons which accept an input, assigns weight to it and sums all of these up. These networks can be trained in order to solve problems more efficiently [Kro08].

The resulting trained model is expected to adequately predict an outcome. This result very important as it ultimately governs the diagnosis, prognosis and chosen therapy for (cancer) patients. Prior to being applied within a clinical setting, however, the stability of the model is required to be assessed. One way to validate is to checking the model with multiple cohorts which are independent from one another. Services such as The Cancer Genome Atlas (TCGA)[1] and The Cancer Imaging Archive (TCIA)[2] offer a wide range of publicly available data which can be used for such validation tasks [RBR+18, vTCTL+20].

[1]TCGA: `https://www.cancer.gov/tcga`
[2]TCIA: `https://www.cancerimagingarchive.net`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

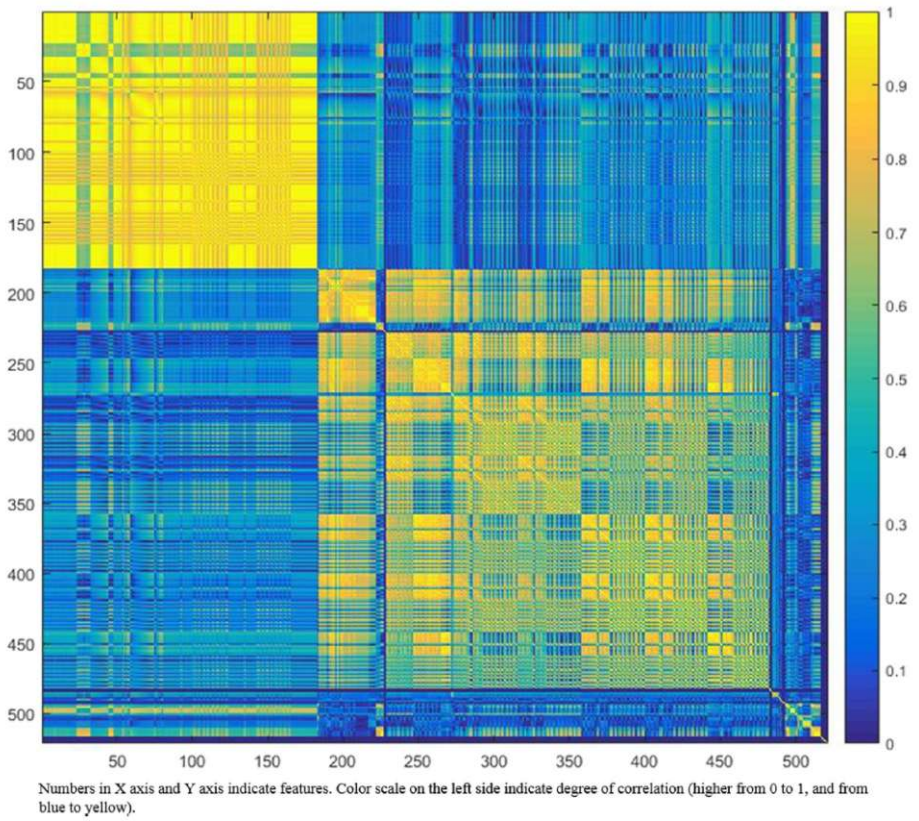


Figure 2.3: Cluster analysis of radiomic features using a heatmap which represents the correlation (ranging from 0 - low to 1 - high) between each pair of radiomic features, with each single-colored block representing one cluster (e.g., yellow block on upper left corner) [RBR+18].

2.2.2 Extracted Radiomic Features

Features extracted from medical images can generally be of qualitative or quantitative nature. Qualitative features are usually linked to semantic descriptions of physical trauma, meaning wounds and lesions. In contrast, quantitative features are derived from clinical images through computational algorithms. These elements include, among others, the shape or texture of tissues. Radiomic features refer to the extracted quantitative features [RPB+16, LvTdJ+17, YA16]. When talking about radiomic features, the quantitative features extracted from medical images is indicated.

Radiomic features can be classified into the following categories: (1) *statistical*, (2) *shape-based*, (3) *model-based*, and (4) *transform-based* [Haj04]. The following subsections describe each feature class. An overview of radiomic features are displayed in Figure 2.4.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

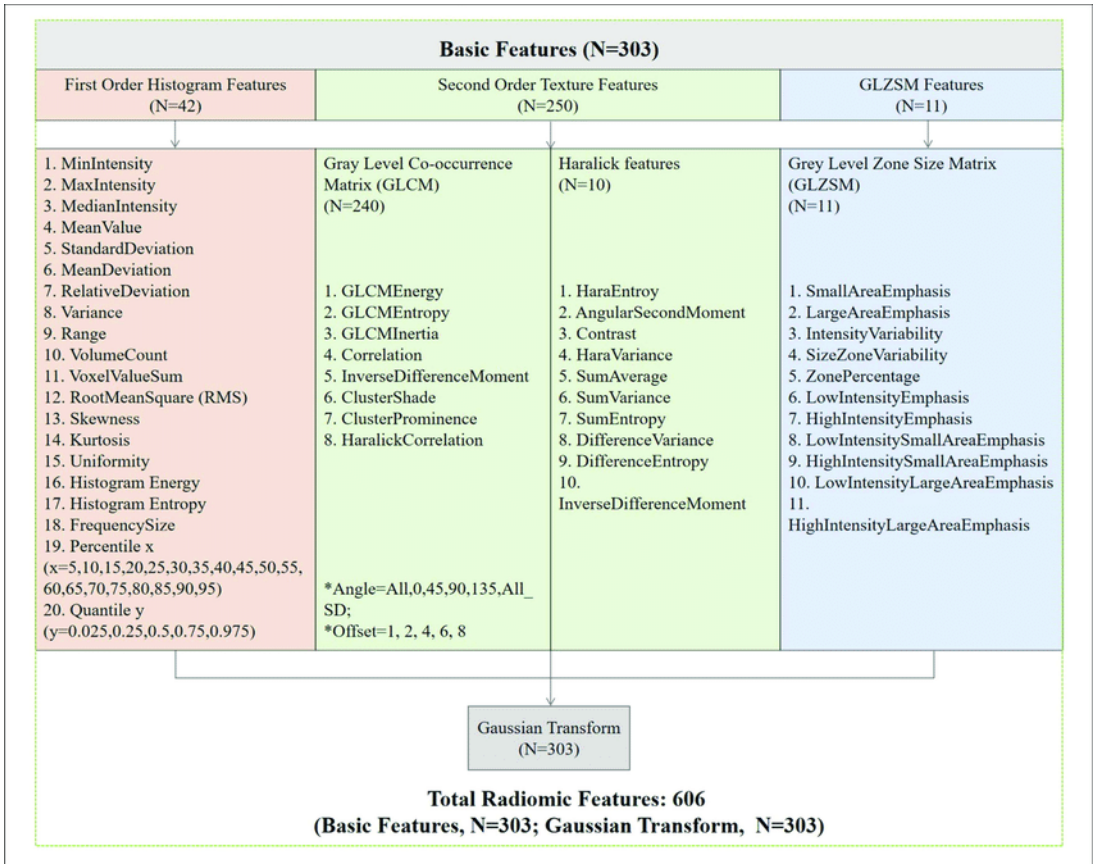


Figure 2.4: Overview of radiomic features [WZZ+19].

Statistical Features

Statistical features rely on the statistical descriptors in an image such as mean, median, and variance of pixel or voxel intensities. This category can be further divided into (1) first-order statistics or histogram-based, (2) second-order statistics or texture-based, and (3) higher-order statistics [MML+20, TMB+18].

First-order statistics or *histogram-based* features define the distribution of individual pixel or voxel values within an image. They rely on characteristics which are based on the statistical measures of an image histogram. These include the simple descriptors mean, median, minimum, maximum, variance, and percentiles of individual pixel or voxel intensities in gray-level images. The more complex measures describe the shape of the intensity distribution of an image. These encompass kurtosis (flatness), skewness (asymmetry), entropy (randomness) and uniformity (energy) [RBR+18, Haj04, ZVA+20].

Second-order statistics or *texture-based* features consider the neighborhood of each pixel and voxel. The focus lies on the spatial relationship of intensities within an image which especially supports exploring (tumor) heterogeneity. This category include the following

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

measures [BKG+14, RBR+18]:

- *Absolute Gradient:* The extent of gray-level intensity of pixels or voxels varying throughout an image [ZVA+20].
- *Gray-Level Coocurrence Matrix (GLCM) or Haralick features*: The spatial arrangements of two neighboring pixels or voxels are described with an gray-level histogram of second-order [Haj04, HSD73, ZVA+20].
- *Gray-Level Run-Length Matrix (GLRLM)*: The spatial distribution of consecutive pixels or voxels in a run with the same gray level depending on the direction are analyzed [Gal75].
- *Gray-Level Size Zone Matrix (GLSZM)*: Similar to GLRLM, but instead of runs, groups or zones of neighboring pixels or voxels are analyzed [TAM14].
- *Gray-Level Distance Zone Matrix (GLDZM)*: As an extension to GLSZM, this method also requires the pixels or voxels to be at the same distance from the edge of the ROI/VOI [TAM14].
- *Neighborhood Gray-Tone Difference Matrix (NGTDM)*: The sum of gray level differences between a pixel or voxel and the mean of the adjacent pixel or voxel is calculated [AK89].
- *Neighborhood Gray-Level Dependence Matrix (NGLDM)*: Similar to NGTDM, but the neighborhood of a pixel is more strictly defined with a predefined distance and dependence criteria [SW83].

Higher-order statistics features are acquired by using filters or mathematical transforms followed by statistical methods. This approach can facilitate the identification of patterns as well as extraction of features from noisy images [RBR+18]. Examples of filters or transforms applied include the Minkowski functionals which supports the analysis of tumor structure [LCK+14] or the Gaussian filter in combination with Laplacian transforms, which aim to suppress noise in images [PS17].

Model Features

Model features focus on the characterization of shapes or objects based on gray-level information in space. They are defined by the estimated parameters which are retrieved from a model computed for the texture generation of a certain ROI/VOI. This so-called auto-regressive model is calculated using the gray levels of four adjacent pixels surrounding one pixel for the gray-level estimation of this pixel [Haj04].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Transform Features

Transform features or wavelet features aim to explore image properties including gray-level patterns within various spaces. These include methods such as the Fourier, Haar, and Gabor transforms which help with the investigation of gray-level patterns. Transform features are particularly valuable for the analysis of frequency patterns and variability within an image [Haj04, MML+20, TMB+18].

Shape Features

Shape features define the shape including its geometric attributes of ROI/VOI. This includes the volume, (maximum) surface, (maximum) diameter, tumor compactness, and sphericity. Specifically the calculation of properties for surface and volume is based on using meshes, which basically are small polygons. Considering compactness and sphericity, the difference between the tumor in the ROI/VOI and a circle (2D) or sphere (3D) is examined [MML+20, ZVA+20, RBR+18].

2.2.3 Applications and Limitations

Radiomics has been applied in various fields including oncology, neurology, and epidemiology. These include applications for providing prognostic models for Multiple Sclerosis lesions [LPL+22] and COVID-19 [SSP+22]. Due to their potential to provide an objective and quantitative assessment of tumor phenotypes, radiomics has particularly been utilized in the field of oncology and personalized medicine. Multiple studies suggest that radiomic features may deliver crucial information for the prognosis of cancer treatment response and outcome. By means of their characteristics they also have been valuable for the identification of tumor tissues based on their malignancy as well as differentiating tumor stages [YA16].

However, radiomic analytics also exhibit technical limitations. These include the susceptibility to varying image acquisitions resulting in large variability among the extracted features. As seen in Figure 2.5, variability concerning tumor segmentation is one exemplary case that may contribute to this issue. Regarding the computations of radiomic features, variations may also appear due to different implementations and general instabilities of these characteristics, which means that the outcome of radiomics analysis will be inconclusive or erroneous [MML+20, YA16].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

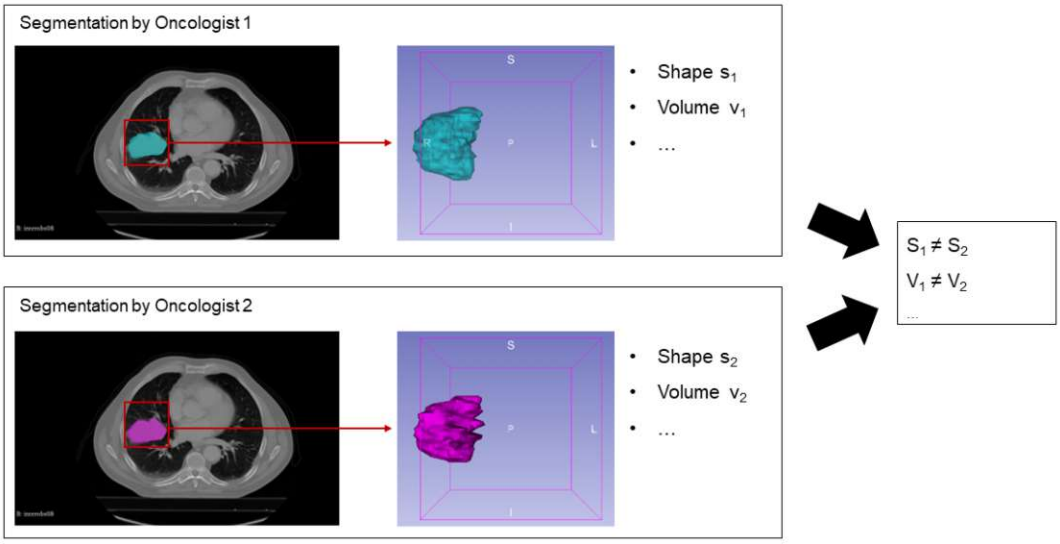


Figure 2.5: Interobserver variability in tumor segmentations resulting in variability among segmented tumors and respective characteristics which again influence the radiomics feature analysis (Input images from Wee et al. [WAKD19] were visualized with 3D Slicer[3]).

[3]3D Slicer: https://www.slicer.org

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 3

Related Work

As probabilistic tumor segmentations and radiomics data are often of complex nature, visual analytics is commonly utilized in order to support the analysis of these. By means of interactive visual representations, the findings of potential patterns within images needs to be facilitated. In this chapter, statistical techniques and visualizations for probabilistic tumor segmentations will be introduced before presenting selected visual analytics tools for radiomics data.

3.1 Probabilistic Tumor Segmentation

Numerous studies resorted to probability approaches to boost the reliability and precision of tumor delineations. By means of statistical techniques, the issue of ambiguity and disagreement due to interobserver variability (Section 2.1) is supposed to be tackled. In this way, information about segmentation estimations and accuracy is gathered.

Naz et al. [NMI10] reviewed *fuzzy clustering*, particularly Fuzzy C-Means (FCM) techniques which investigate the degree of pixels or voxels affiliated with certain segments. They concluded that FCM produced promising results, provided that the conventional algorithm was altered using spatial data. These results were also produced in images with artifacts (e.g., noise), meaning that their approach is considered robust against image artifacts. However, this advantage may only apply to MRI images, as Naz et al. did not investigate the effectiveness of their technique in other medical imaging types such as CT in their research.

Lelandais et al. [LGM+14] created the Estimation of Imperfect Information (EVEII) method which is based on the Belief Function Theory (BFT) and uses spatial information. The certainty of voxels believed to reflect tumor tissue is represented by means of color intensities. Similar to the approach by Naz et al., their proposed method was also robust against noise in images. When it comes to images with large quantity of noise, it even

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

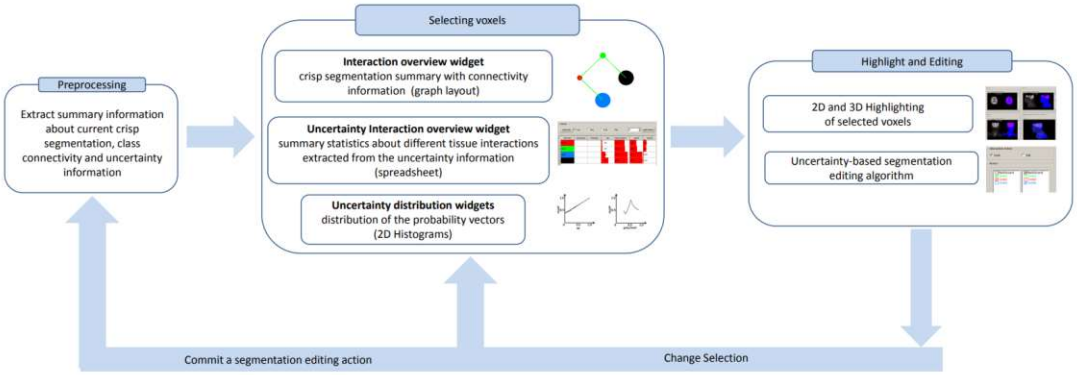


Figure 3.1: The three steps of the ProbExplorer by Saad et al. [SMH10], consisting of preprocessing, voxel selection, and highlighting and editing for the analysis of probabilistic segmentation.

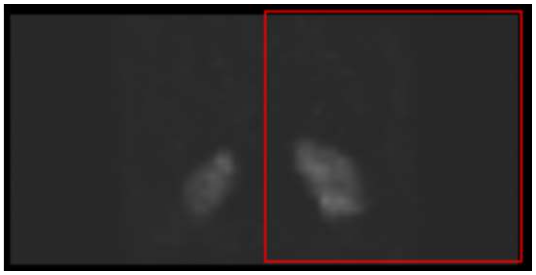
outperformed other approaches using FCM. As Lelandais et al. only considered a smaller database, the exact robustness of their method has yet to be assessed with larger input data.

Saad et al. [SMH10] developed an interactive visualization tool for analyzing probabilistic segmentation in clinical images named ProbExplorer, incorporating the aspect of user engagement. The framework involves three steps: (1) preprocessing including the extraction of quantities, (2) voxel selection implicating the specification of ROI/VOI, and (3) highlighting and editing comprising the application of specific actions on the chosen voxels. For the second step, ProbExplorer provides the following three widgets presenting: (1) an interaction overview showing a segmentation summary and its connectivity in a graph layout, (2) an uncertainty interaction overview displaying a statistical summary about tissue interactions in a spreadsheet, and (3) uncertainty distribution presenting the distribution of probability vectors in 2D histograms. Following the ROI/VOI selection, distinct colors are used for each class (e.g., left kidney, right kidney). In order to provide context, the surrounding tissues are presented in gray levels. With ProbExplorer, Saad et al. devised a novel way of displaying segmentation probability by means of visualizing information about probability and semantics on interactive widgets. Examples of their designed visualizations are illustrated in Figure 3.2, which shows results from a case study on abnormal renal behavior in the left kidney.

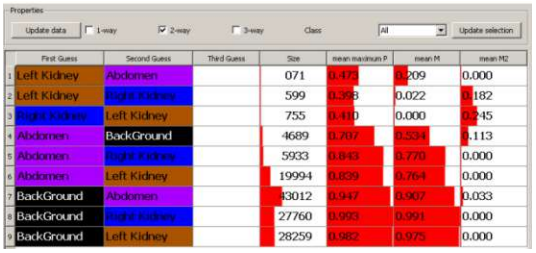
Due to latest progresses and successes of artificial intelligence in medical image segmentation, more recent studies followed deep learning techniques. The most popular architecture used are U-Nets developed by Ronneberger et al. [RFB15]. It is a convolutional network specifically developed for biomedical image segmentation and has shown great performance in delivering both fast and precise segmentation [RFB15].

Chotzoglou and Kainz [CK19] explored the relationship between interobserver variability and deep network segmentation ambiguities. Using U-Nets, they have found correlations

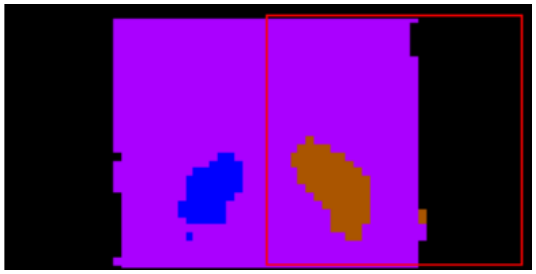
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



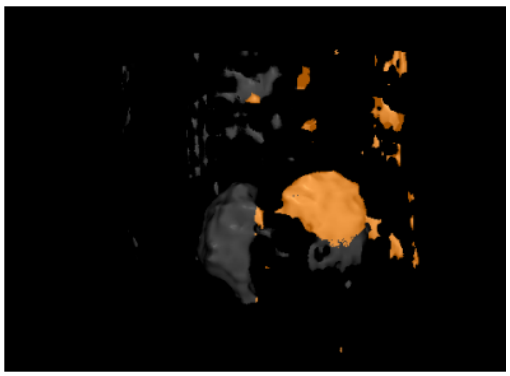
(a) Coronal/frontal slice of kidney.



(b) Uncertainty interaction overview widget.



(c) Selection of ROI over left kidney.



(d) Visualization of highlighted voxels which are labeled left kidney.

Figure 3.2: Widgets for case study on abnormal renal behavior in the left kidney with ProbExplorer [SMH10].

between annotator and models in delineation uncertainty and stress the benefit of reevaluating uncertain segments manually after the automatic delineation. Similar findings were made by Chlebus et al. [CMT+19], who aimed to reduce interobserver variability by means of manual corrections of automatic segmentations based on convolutional neural networks (CNN) with a U-Net like architecture. Their approach reduced the time of image delineation significantly without compromising the quality, as the resulting segmentations were comparable to those obtained manually.

Jungo et al. [JME+18] also investigated these fusion techniques, but with the observers delineating the tumors first before combining these inputs for the automatic generation of labels. In other words, they used multi-observer annotations to train deep-learning models for estimating parameter uncertainty within clinical images. In doing so, they have explored different fusion techniques of input such as the intersection of the labels or inclusion all labels (no fusion). They concluded the profit of using fusion methods to be highly dependent on the segmentation precision of observers, regardless of the type of fusion technique itself. Figure 3.3 shows the comparison of uncertainty estimates resulting from five different fusion techniques.

Considering our aims and the existing literature, we plan to evaluate the effect of

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

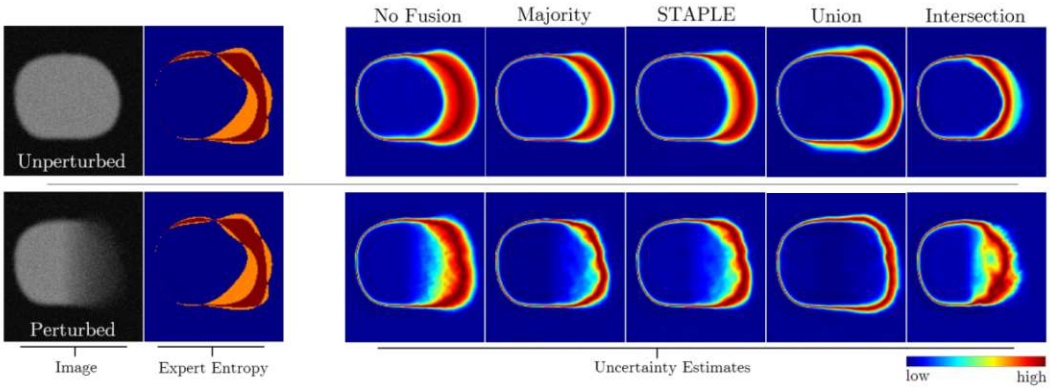


Figure 3.3: Uncertainty estimations using five fusion techniques (no fusion, majority voting, STAPLE, union of segmentations by all observers, intersection) on both unperturbed and perturbed images [JME+18].

probabilistic segmentation on the radiomics analysis. As the calculation of segmentation probabilities is not the main focus of our thesis, we will consider less complex estimation algorithms which are not based on deep learning approaches. Instead, we will focus on observing generating segmentation probability thresholds based on interobserver variability present in the data set used [WAKD19]. We subsequently explore the effect of different segmentation accuracies on our radiomics analysis, which is the main goal for our research.

3.2 Visual Analytics of Radiomics Data

Radiomics data often comprise large amounts of information, making its analysis particularly challenging [MWLH+20]. Especially the high dimensionality of the derived features aggravate the process of understanding data and finding underlying patterns [MML+20]. Even though radiomics does improve the analysis of clinical imaging data, the extracted information has to be visualized properly so that underlying patterns may be unraveled, contributing to predicting patient outcome [MWLH+20, YJY+17].

Previous studies have presented visual analytics tools for observing radiomic features. Most of these are interactive, making it possible to explore relationships between features more freely and eventually revealing correlations with clinical data. The multiple views available within a tool also allow the exploration of different types of data obtained from patient cohorts. The following subsections present visualizations found in visual analytics tools for radiomic features. They will be introduced based on type of task, which are (1) feature distribution (Subsection 3.2.1), (2) feature correlation (Subsection 3.2.2), (3) cluster analysis (Subsection 3.2.3), and (4) interaction (Subsection 3.2.4). As the field of radiomics is relatively new, there were no specific criteria for the selection of papers.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

3.2.1 Feature Distribution

The distribution of radiomic features is helpful for the verification of found patterns within the imaging data. The idea is to provide an overview of relationships between quantitative features. Due to the numerical characteristic of this task, prior research have restorted to statistical chart visualizations such as bar charts [YJY+17].

Yu et al. [YJY+17] used frequency plots for visualizing the distribution of radiomics features among a patient cohort (Figure 3.4). Based on the extracted feature, the scaling of the chart axes is adjusted. For example, the distribution of the *Volume* feature is presented with a volume size range (x-axis) in relation to the number of patients exhibiting the volume size (y-axis). By means of the the curve connecting the respective values as well as the blue-colored area below the curve, Yu et al. provided a clear overview of the frequency distribution of each feature regarding its value.

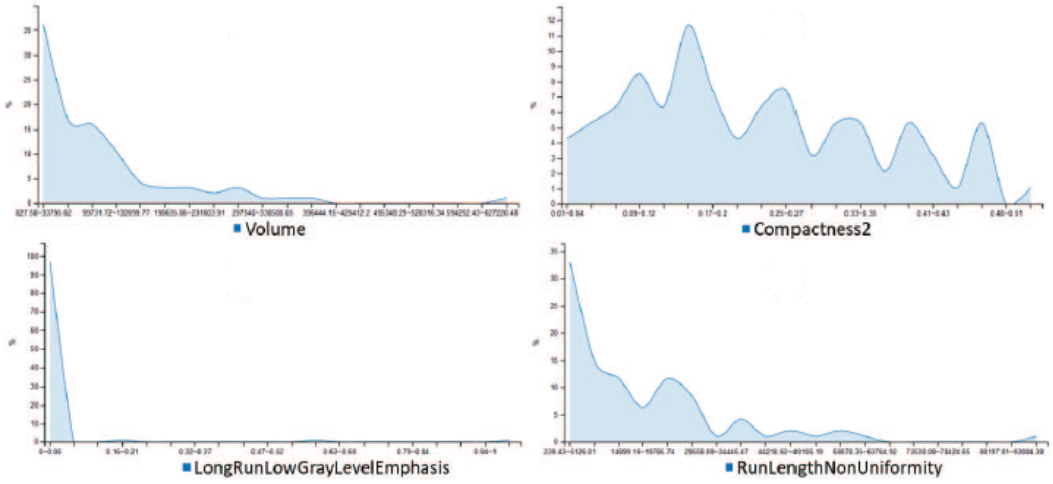


Figure 3.4: Feature distribution of the features *Volume* (top left), *Compactness* (top right), *Long Run Low Gray Level Emphasis* (bottom left), and *Run Length Non Uniformity* (bottom right) showing the frequency of certain values across all patients [YJY+17].

In their visual analytics framework called *AnaFe*, Gutenko et al. [GDKB17] utilized feature histogram charts for visualizing the distribution of features (Figure 3.5). Similar to the frequency plot described before, the horizontal axis (x-axis) represents ranges of the measurement of the feature (e.g., length), whereas the vertical axis (y-axis) describes the number of subjects exhibiting the specific range of measurement. Each bin or segmented column consists of stacked rectangles which can be colored in four different color schemes: (1) age (shades of violet), (2) gender (blue for “male” and pink for “female”), (3) disease status (green for “healthy” and violet for “sick”), or (4) unique subject (one color per individual subject). As Gutenko et al. also considered the feature progression of the subjects over time, each feature comes with two visualizations. Whereas one is presenting

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

the distribution of absolute measurement recorded at a specific time, the other shows the relative changes within a period of time.

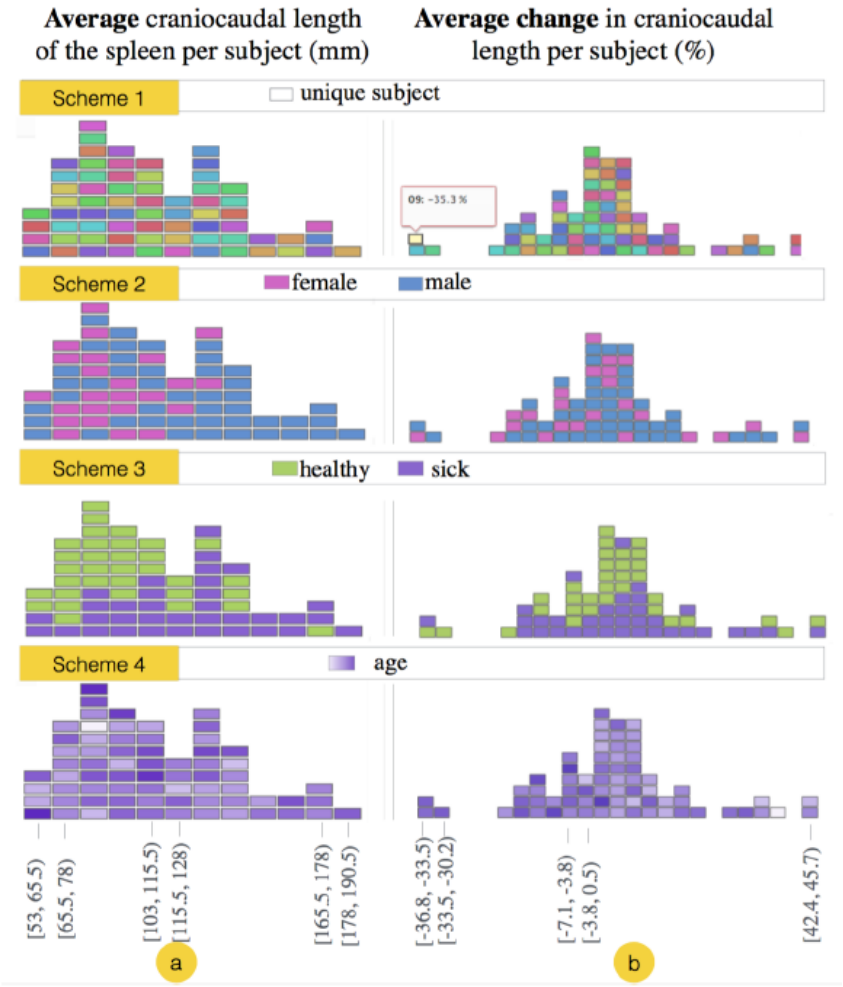


Figure 3.5: Feature histogram charts for the *average craniocaudal length* and *average change in craniocaudal length per subject* in four color schemes (unique subject, biological sex (female/male), health status (healthy/sick), and age), respectively [GDKB17].

3.2.2 Feature Correlation

Investigating the relationship between features is crucial for building prediction models. Information about which features correlate with one another allow the deduction of dependencies between two or more dimensions. Based on the number of features which are set into correlation, past studies have utilized visualizations of different complexity [YJY+17]. The following subsections present representations categorized by number of

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

features considered, namely pairwise correlation (2D) and multidimensional correlation (going beyond plain 2D).

Pairwise Correlation

When investigating the pairwise correlation among features, the focus is set to analyzing the linear relationship between two features respectively. Previous studies fulfilled this task by providing an overview of pairwise in one total representation or producing one visualization for each feature pair. While the first typically consists of a matrix organization with color encoding, the latter usually takes the form of a scatterplot. The following examples show visualizations of both types.

Raidou et al. [RvdHD+15] explored the pairwise correlations among features by means of a *scatterplot matrix (SPLOM)* (Figure 3.6). This visualization type provides an overview of correlations between two features through a matrix organization. In order to further simplify the SPLOM, Raidou et al. calculated the Pearson correlation for each pair of features. The *Pearson correlation* or Pearson ρ, is a measure of linear relationship between two variables, ranging from -1 (total negative or inverse correlation) to 1 (total positive correlation), whereas 0 means that there is no correlation [WHL+20]. The rectangles in the SPLOM are colored in a red-to-blue color scheme. So for example, if one rectangle is colored dark blue, the respective features X and Y are positively correlated with one another, meaning that both variables change in the same direction.

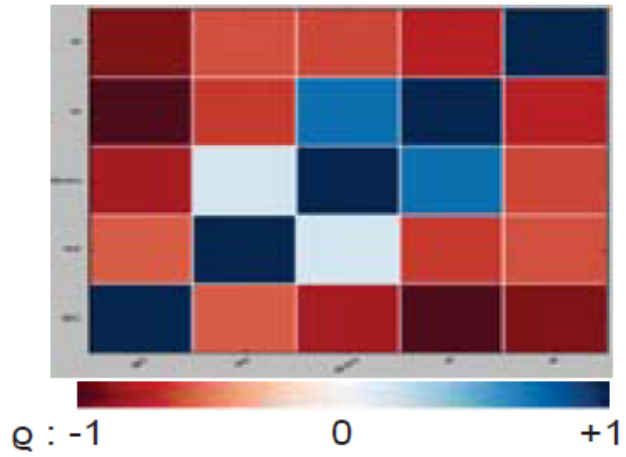
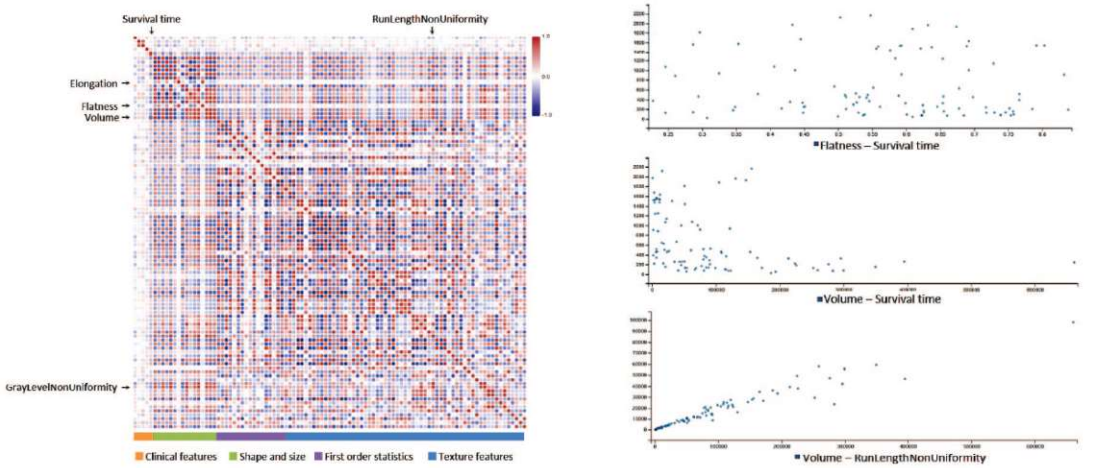


Figure 3.6: Simplified SPLOM in a red-to-blue color scheme [RvdHD+15].

As seen in Figure 3.7, Yu et al. [YJY+17] focused on providing more details by using *correlation matrices* (Figure 3.7a) as well as *scatterplots* (Figure 3.7b). They considered radiomic features, or more specifically, shape and size, first-order statistics, and texture features, as well as clinical features. Similar to the work of Raidou et al. [RvdHD+15], they also use a red-to-blue color map in the correlations matrix to show the calculated

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Pearson correlation, but with an inverted scale. This means that in this representation, red implies that there is a positive correlation ($\rho = +1$), whereas blue signifies a negative correlation ($\rho = -1$). In addition to the correlation matrix, Yu et al. also provided scatterplots for each pair of features. These show a more detailed view by displaying the relationship between two variables of each data subject through a point. Based on the distribution of the data points, the strength of correlation between the features can be derived.



(a) Correlation matrix showing the Pearson correlation between pairs of features.

(b) Scatterplots showing the correlation between selected feature pairs.

Figure 3.7: Two visualization types for representing pairwise radiomic feature correlation [YJY+17].

As our thesis aims to investigate the effect of segmentation probability on the radiomics analysis, we do not intend to particularly focus on the correlation of feature pairs. We anticipate the interobserver variability affecting the radiomic feature calculation significantly, hence producing different results for each segmentation conducted by one observer. As a result, we shift the focus from solely investigating feature correlations to the variability of feature values based on interobserver variability and the respective segmentation probability.

Multidimensional Correlation

In contrast to pairwise correlations, multidimensional correlations consider more than two features in the relationship analysis. When analyzing multidimensional correlations, the degree of relation between features as well as their interdependency are explored. Due to more dimensions being involved, the respective visualizations for multidimensional correlations analysis also become more complex. For this task, several studies employed *parallel coordinate plots (PCP)* [Ins85]. This visualization type is very commonly used for this purpose as it facilitates the detection of high-dimensional patterns. It consists of

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

several vertical axes, whereas each represents a variable, or in this specific case, features. The values for the variables are represented as polylines, which are lines connected across each axis. In the context of radiomics, a polyline represents the values captured from one patient. It should be noted that the design of the PCP plays an essential role in the quality of analysis information acquired. Consequently, it is crucial to adequately scale the feature axes as well as consider the order of these when producing PCPs. This is due to feature dimensions visually further apart and with not standardized scaling dimensions being hard to compare [ID90, MWLH+20, RvdHD+15].

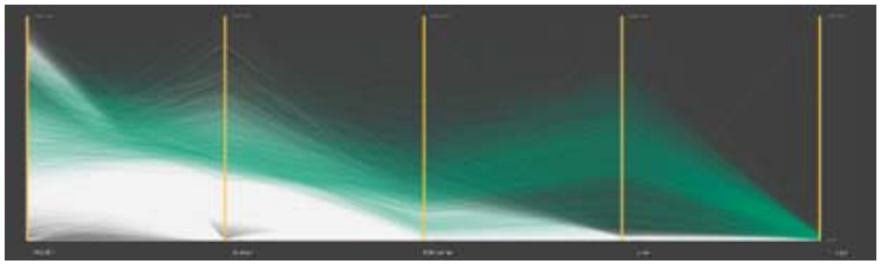
A number of authors have produced PCPs for analyzing clinical and radiomic features (Figure 3.8). Color schemes were also considered in the representations to highlight interesting patterns within the plots. Raidou et al. [RvdHD+15] employed PCPs for exploring intra-tumor tissue characterizations (Figure 3.8a). For readability reasons, they rendered the polylines with a low opacity and used a color map consisting of white, green and black to further mark clusters. Yu et al. [YJY+17] also used colors to denote clusters in their PCPs but with sharply rendered polylines and a red-and-blue color scheme (Figure 3.8b). To further assist the correlation analysis process, they considered brushing approaches. These allow closer investigations on correlation trends between two dimensions by filtering polylines based on their values in a dimension. So in other words, "brushing" an axes signifies means to only focus on data which lie within a defined value range in that dimension. Yu et al. mark the brushed area by means of a green semi-opaque rectangle on the axes. As seen in Figure 3.8b, the negative correlation is more noticeable after brushing. Similarly, Moerth et al. [MWLH+20] visualized sharp polylines and incorporated brushing techniques. In their visualization, the brushed axes are marked with red rectangles which may be connected in case two axes were brushed. Accordingly, the filtered polylines which lie within the selected ranges are colored blue, whereas the other lines are colored grey. This color scheme is also utilized in the corresponding scatterplot which Moerth et al. provided in addition to the PCP. In the scatterplot, each patient is represented by a data point containing a small glyph which shows the tumor characteristics shape and size.

3.2.3 Cluster Analysis

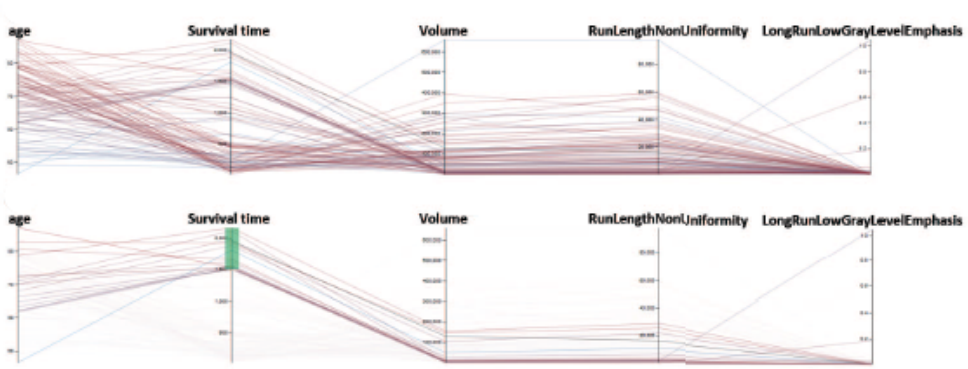
The analysis of clusters plays an important part in revealing patterns within data [PLG+15]. It supports the identification of both significant and stable radiomic features for prognostic prediction as well as the discovery of groups which exhibit similar characteristics [YJY+17]. Past research utilized various types of visualizations including dendrograms and heatmaps to clearly represent cluster analysis data.

In their visual analytics tool iVAR, Yu et al. [YJY+17] utilized heatmaps (Figure 3.9) for the exploration of feature and patient clusters. Their clustering results are obtained using unsupervised hierarchical clustering, which is a method that produces partitions of data based on hierarchy [Hal18]. The map conveys information about the feature values through color, using a green-to-black-to-red color scheme. The cluster information is communicated by means of the two dimensions, with the vertical axis showing patient

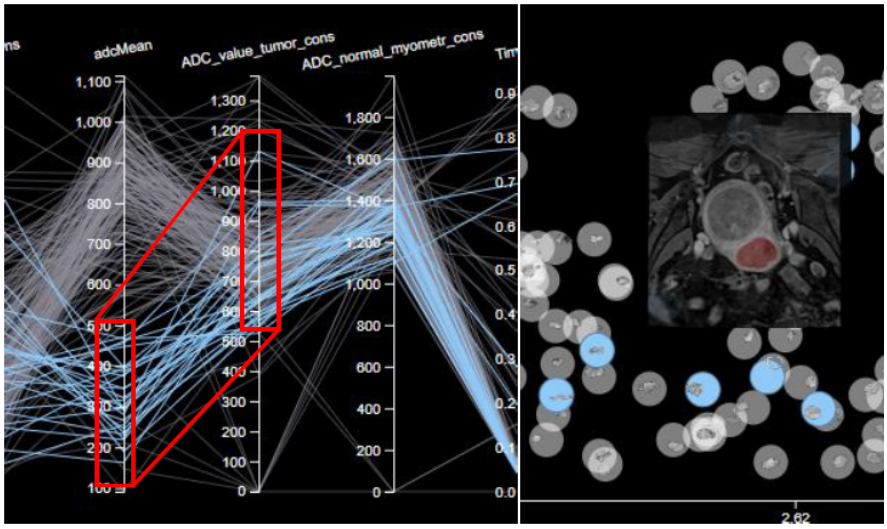
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) PCP with low opacity and a white, green, and black color scheme [RvdHD+15].



(b) PCP with a red and blue color scheme without (top) and with brushing feature (bottom) [YJY+17].



(c) PCP and scatterplot in blue and grey color scheme with brushing feature [MWLH+20].

Figure 3.8: Overview of PCPs used for investigating multidimensional correlations in radiomics data.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

clusters and the horizontal axis representing feature clusters. In this way, Yu et al. support a simultaneous cluster analysis of patients and also features, both in terms of similar feature values.

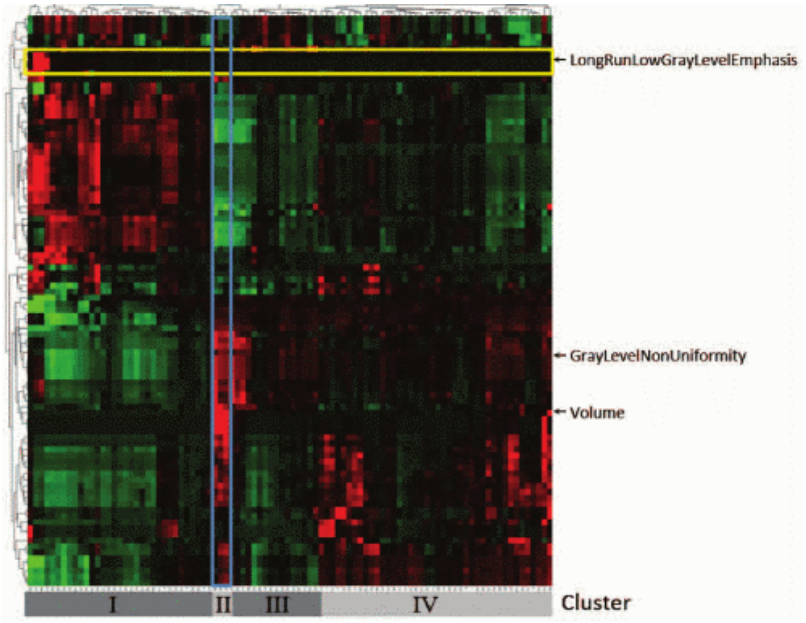


Figure 3.9: Heatmap showing cluster information about patients through the vertical axis (e.g., blue box) and features through the horizontal axis (by means of color/values, e.g., yellow box), using a green-to-black-to-red color scheme to represent the feature values [YJY+17].

Moerth et al. [MEH+22] visualized hierarchical cluster data by means of a dendrogram (Figure 3.10) in their tool ICEvis. The visualization represents hierarchical structure of clusters, with the links between subclusters determining the distance between each other. It also includes a dashed horizontal line, which does not represent information about the cluster information per se, but indicates the number of clusters selected in their visual analytics tool.

Raidou et al. [RvdHD+15] revealed clustering data beyond hierarchical structure and cluster distribution by providing metrics about the validity of clusters. By this, they aimed to facilitate the comparison between clusters in terms of their internal structure. Thereby, they used the commonly-used metrics *cohesion (WSS)*, which represents the relation within one cluster, *separation (BSS)*, which shows the distinction between clusters, and *average silhouette coefficient (s)*, which represents the validity of a cluster regarding cohesion and separation [TSK05]:

Cohesion (WSS) is defined as

$$WSS = \sum_{x \in C} (x - m)^2 \tag{3.1}$$

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

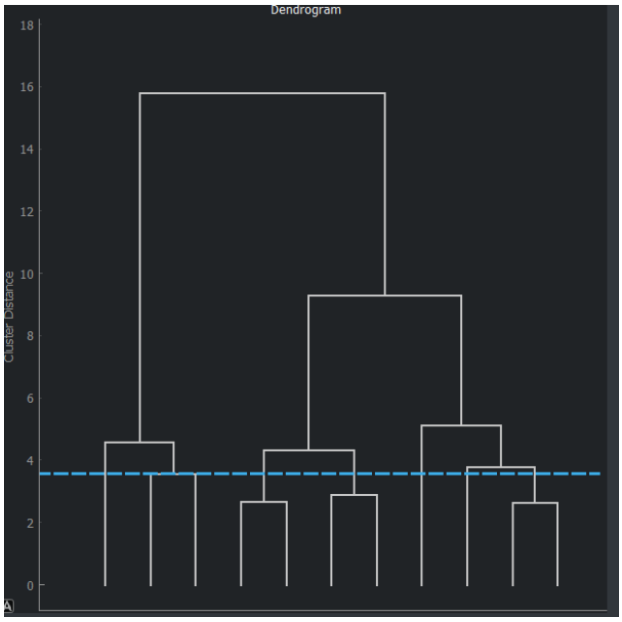


Figure 3.10: Dendrogram representing hierarchical clustering and the distance between clusters through the length of vertical lines [MEH+22].

whereas C reflects one cluster, x the vector of feature values, and m the vector of mean feature values within one cluster.

Separation (BSS) is defined as

$$BSS = \sum_i |C_i|(m - m_i)^2 \tag{3.2}$$

whereas C_i reflects the selected clusters, i the respective size of the clusters, m the vector of overall mean feature values, and m_i the vector of mean feature values for the selected clusters.

Average silhouette coefficient (s) is defined as

$$s = \frac{BSS - WSS}{max(BSS, WSS)} \tag{3.3}$$

which ranges between 0 and 1 and can usually be defined in categories (e.g., 0 – 0.25: bad-defined, 0.26 – 0.5: weak, 0.51 – 0.75: reasonable, and 0.76 – 1: excellent).

The results were visualized by means of a sphere for each cluster (Figure 3.11). The appearance of the sphere is dependent on the metrics, with cohesion being reflected through size, separation through distance, and silhouette coefficient through color shade. Furthermore, each sphere is filled with a different color to make the clusters easily distinguishable. Ultimately, the approach by Raidou et al. enables a thorough cluster analysis by giving visual insights about the cluster distribution as well as their validity

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

in an abstracted manner. However, this abstraction does not scale well if many clusters are present in the data.

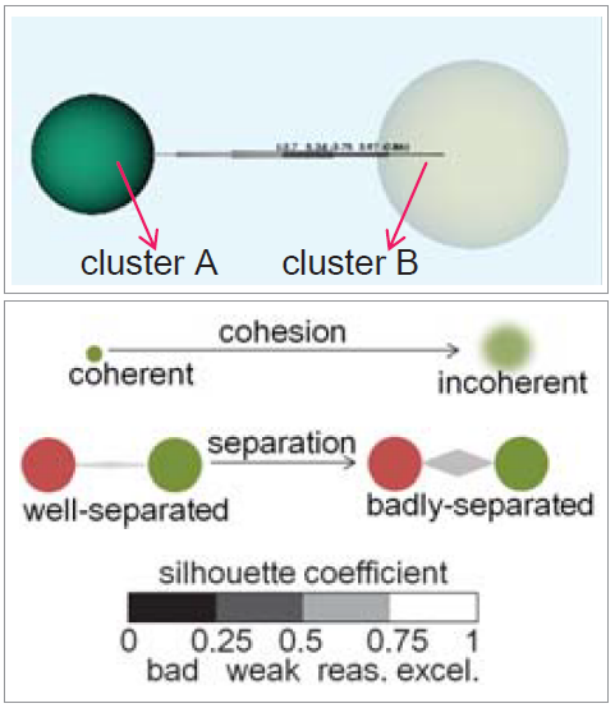


Figure 3.11: Cluster analysis using spheres to represent clusters, whereas the size, distance, and color shade is dependent on the metrics *cohesion (WSS)*, *separation (BSS)*, and *average silhouette coefficient (s)* [RvdHD+15].

We strive for providing a thorough but also simple cluster analysis in our thesis. Therefore, we will follow the approach presented by Raidou et al. [RvdHD+15] in order to obtain information about the validity of clusters through calculated metrics. We will then visualize these results using dendrograms, as employed by Moerth et al. [MEH+22]. In this way, we show the hierarchical structure of clusters as well as the validity of these through one visual representation.

3.2.4 Interaction

Recent studies increasingly provided interactivity within their visual analytics tools. This is due to radiomics data usually being complex, which make it hard to analyze patterns and therefore predict clinical outcome. The possibility to customize the information displayed allows an optimized exploration of radiomic features and their relationships with each other. In this section, the interaction aspects found in the previously presented visual analytics tools are elucidated.

In their visual analytics tool called *AnaFe*, Gutenko et al. [GDKB17] provided the possibility to zoom, filter, and customize query parameters as interaction features. In this

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

way, they allow the user to visually hide data not significant for the analysis, avoiding information overload on visualizations. Furthermore, they have also provided clarity through only presenting information when users hover over certain elements in the representations. For example, the rectangles in the aforementioned histogram chart (Figure 3.5) exhibit the feature to show additional information about the respective value range when hovering over them.

Both Yu et al. [YJY+17] and Moerth et al. [MWLH+20] assured interactivity in their visual analytics tools through multiple interactive linked panels [YJY+17]. So when filtering or brushing data (Subsection 3.2.2) in one visualization, the connected views are updated accordingly. The brushed area specifying the ranges on an axis can be also be adjusted anytime by moving the corresponding rectangle. This interaction feature allows the users to customize the panels visualizations interactively, aiding them in their exploration and formulation of hypotheses. Figure 3.12 shows examples for brushing and the corresponding results on PCPs compared to the original result.

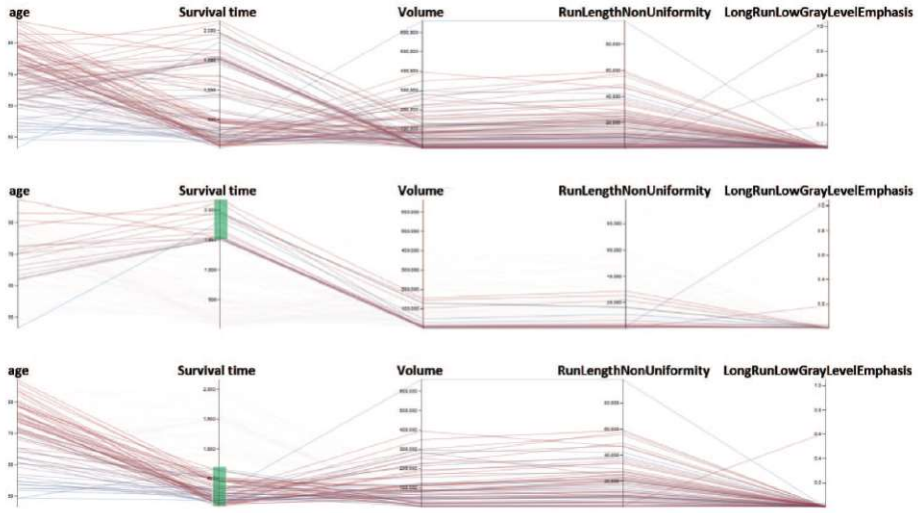


Figure 3.12: PCPs for five features showing the original plot (top) and the plot after brushing (middle and bottom) by means of a green rectangle on the "Survival time" feature axis [YJY+17].

For the visual analytics part of our thesis, we plan to provide visualizations for analyzing the distribution and correlation of radiomic features and clinical data. Similar to the planned visualizations for the probabilistic segmentation task, we also intend to use color schemes and interaction possibilities to strengthen user engagement. We will link our views with each other similar to the tools designed by [YJY+17] and [MWLH+20]. As we explore the effects of probabilistic segmentations on radiomic features, the visualizations for both the segmentation and the radiomics part will be linked. This ensures the possibility to directly explore the influences of different tumor delineations on the radiomics data.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

3.3 Summary

The literature review shows several studies which have been published in the past years. The studies on probabilistic tumor segmentation comprised simple techniques such as FCM but also more complex algorithms involving deep learning mechanisms using U-Nets. Current research on visual analytics of radiomics data show different visualization designs for similar tasks, which include visualizing feature distribution, feature correlation and cluster analysis. Most of these also exhibit interaction possibilities which support customization and, related to this, also the analysis process. In Table 3.1, the reviewed papers are summarized and compared to meeting the aims of the thesis. As mentioned in Section 3.1, the reviewed papers on probabilistic segmentation based on deep learning (e.g., [CMT+19, CK19, JME+18]) are not considered for this thesis and therefore are not be included in the table.

Despite various research published on these topics, previous studies have almost exclusively focused on either solely visualizing probabilistic segmentations or analyzing radiomic features, but not the influence between both tasks. To our knowledge, no prior studies aimed to examine this topic. Thus, we focus on exploring the impact of interobserver variability in tumor segmentations on the visual analytics of radiomic features.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Table 3.1: Overview of reviewed literature compared to aims of thesis: A (Visualization of radiomics with regard to clinical data), B (Investigation of the effect of probabilistic tumor segmentation on the radiomics analysis), and C (User interaction).

Literature (Author, Year)	Subject	Aims Fulfilled (Section 1.1)		
Naz et al., 2010 [NMI10]	Probabilistic Segmentation	☐ A	☑ B	☐ C
Saad et al., 2010 [SMH10]	Probabilistic Segmentation	☐ A	☑ B	☐ C
Lelandais et al., 2014 [LGM+14]	Probabilistic Segmentation	☐ A	☑ B	☐ C
Raidou et al., 2015 [RvdHD+15]	Visual Analytics on Radiomics	☑ A	☐ B	☐ C
Gutenko et al., 2017 [GDKB17]	Visual Analytics on Radiomics	☑ A	☐ B	☐ C
Yu et al., 2017 [YJY+17]	Visual Analytics on Radiomics	☑ A	☐ B	☑ C
Moerth et al., 2020 [MWLH+20]	Visual Analytics on Radiomics	☑ A	☐ B	☑ C
Moerth et al., 2022 [MEH+22]	Visual Analytics on Radiomics	☑ A	☐ B	☐ C

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 4

Methodology

The focus of this thesis is the development of the visual analytics tool *ProSeRa* which investigates the effect of tumor segmentation on the radiomics analysis. We thereby aim to provide visualizations for the analysis of radiomics with respect to clinical data (**Aim A**), explore the radiomics analysis with regard to probabilistic segmentation (**Aim B**), and enable user interaction in *ProSeRa* for thorough analysis possibilities (**Aim C**). In order to reach these aims, we searched for appropriate data sets, preprocessed the data, and designed and implemented visual interfaces. This chapter elucidates the used materials, the design as well as the implementation of *ProSeRa*.

4.1 Materials

Before starting with the implementation of *ProSeRa*, we searched for appropriate data sets for investigating the effect of probabilistic tumor segmentation on the radiomics analysis. Based on aims **A** (Visualize and analyze radiomic features together with respective clinical data), and **B** (Investigate the effect of probabilistic tumor segmentations on the radiomic features analysis), we set the following search criteria for a data set: The desired data set should (1) contain medical images of several cancer patients from which we can extract radiomics features, (2) comprise clinical information (e.g, histology) to assess our results based on ground truth, and (3) include tumor segmentations by different observers or algorithms to explore interobserver variability and the corresponding segmentation probability.

Based on our search criteria, we chose the *NSCLC-Radiomics-Interobserver1* data set [WAKD19], which was published in The Cancer Imaging Archive (TCIA)[1]. This collection consists of Computed Tomography (CT) images from 22 non-small cell lung cancer (NSCLC) radiotherapy patients, respective tumor segmentation data by five radiation

[1]TCIA: `https://www.cancerimagingarchive.net`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

oncologists, and clinical patient data. In the following subsections, both data types, namely imaging data including segmentations and clinical data are described in more detail.

Image and Segmentation Data

The *NSCLC-Radiomics-Interobserver1* data set consists of CT images from 22 NSCLC radiotherapy patients prior to treatment. The CT data for one patient contains roughly 154 to 178 CT slices, with each having a resolution of 512×512 pixels. In the data set, the patients are referred to as *interobsXX*, whereas *XX* stands for a patient number, which is not in order and ranges from 5 to 34. In addition to the medical images, the collection comprises tumor segmentation data for 21 out of 22 patients. This includes manual delineations as well as semi-automatic delineations by five radiation oncologists. The latter was performed with an automatic tool followed by manual adaptions of the segmentation by the same observers.

All image and segmentation data were provided as Digital Imaging and Communications in Medicine (DICOM) files. The segmentation data was extracted into two DICOM modalities, namely RTSTRUCT and SEGMENTATION. Whereas DICOM RTSTRUCT consists of contour points of the tumor outline, DICOM SEGMENTATION comprises binary masks of the tumor as delineation information.

The segmentation data comprises numeral tags which denote information about the tumor borders, the type of segmentation (manual or semi-automatic) and the radiation oncologist performing the segmentation. We used the tool 3D Slicer[2] to visually observe the tumor delineations within an anatomic context. An example is illustrated in Figure 4.1. The following taxonomy was established [WAKD19]:

- **"GTV-1"**: Index tumor, particularly the gross tumor volume (GTV)
- **"vis"**: Manual segmentation by radiation oncologists
- **"auto"**: Semi-automatic segmentation by an automatic tool followed by manual edits by radiation oncologists
- **"1" ... "5"**: Individual radiation oncologist

Clinical Data

Aside from the imaging and segmentation data, the *NSCLC-Radiomics-Interobserver1* collection comprises clinical patient data which was provided in a comma-separated values (CSV) file. For each of the subjects, information about sex, age, histology (adenocarcinoma, NSCLC, large cell carcinoma, squamous cell carcinoma, or undifferentiated lung carcinoma), tumor location, clinical T-stage, clinical N-stage, clinical M-stage, and overall stage are disclosed. An excerpt of the CSV file is presented in Figure 4.2 .

[2] 3D Slicer: `https://www.slicer.org`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

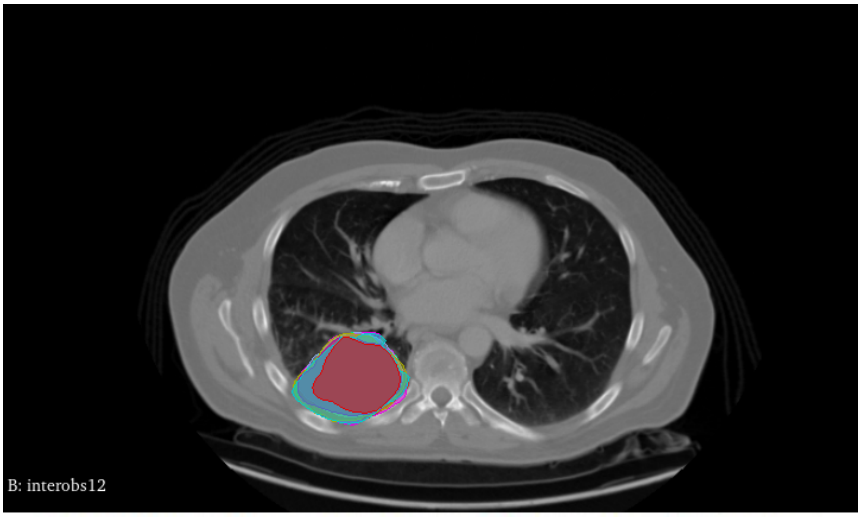


Figure 4.1: Axial view of `interobs12` on slice 107 showing the tumor segmentations (top) and the respective taxonomy (bottom).

PatientID	Sex	Age.years	Histology	TumourLocation	Clinical.T.Stage	Clinical.N.Stage	Clinical.M.Stage	Overall.Stage
interobs05	male	65	adenocarcinoma	right.middle.lobe	T2	N2	M0	IIIa
interobs06	male	82	squamous.cell.carcinoma	left.hilum	T1	N2	M0	IIIa
interobs08	male	66	adenocarcinoma	right.middle.lobe	T2	N2	M0	IIIa
interobs10	female	57	large.cell.carcinoma	left.upper.lobe	T1	N2	M0	IIIa
interobs11	male	74	adenocarcinoma	right.middle.lobe	T2	N2	M0	IIIa
interobs12	male	50	adenocarcinoma	right.lower.lobe	T2	N2	M0	IIIa
interobs13	female	68	adenocarcinoma	left.lower.lobe	T2	N0	M0	Ib
interobs14	male	77	undifferentiated.lung.carcinoma	right.lower.lobe	T2	N2	M0	IIIa

Figure 4.2: Excerpt of clinical data from [WAKD19].

4.2 Design of ProSeRa

Our visual analytics tool *ProSeRa* is a web application. We have opted for this due to the availability of various open source frameworks such as Angular[3] for creating web applications, which provide compatibility to libraries such as D3.js[4], which we use for designing the visualizations. These frameworks also enable us to create modern and contemporary web applications with user interface (UI) component toolkits such as Angular Material[5] and Bootstrap[6], which we both use for designing our web application.

[3] Angular: `https://angular.io`
[4] D3.js: `https://d3js.org`
[5] Angular Material: `https://material.angular.io`
[6] Bootstrap: `https://getbootstrap.com`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

In order to assure that all aims (Section 1.1) are fulfilled, we design our web application involving all planned visualizations and user interactions. This includes wireframes of the user interface as well the interactive aspects of these. Based on the aims of the thesis, we state the contributions to the functionalities and the intentions behind the chosen visual representations.

4.2.1 Design of User Interface

As already stated in Section 1.1, we aim to create an interactive tool which visualizes radiomics data to empower clinical researchers to explore the effect of probabilistic tumor segmentation on the radiomics analysis. We pursue these aims by providing a web application containing several visualizations which are interconnected with each other. All information are displayed on one page in order to enable a comprehensive analysis of various aspects of our data without navigating to other web pages. In order to ensure a clear presentation of data with such high complexity, we split the information based on analysis focus (e.g., cluster analysis) on different components.

We pursue **Aim A** by planning visualizations which represent radiomics data with respect to clinical data. Our goal is to provide a thorough radiomics analysis to the user, hence planning two components which focus on different aspects of our results. One component concerns the correlation of radiomics data including the details about the patient cohort (Figure 4.3), whereas the other component focuses on the cluster analysis (Figure 4.4).

Figure 4.3 shows the wireframe for the UI component which primarily provides an overall insight into the potential correlations and clusters of radiomics with clinical data. For the sake of clarity, we plan two separate visualizations: One for showing the correlation between radiomics data (Annotation 1) and another for representing statistical information about the patient cohort (Annotation 2). By this, we enable the investigation of potential patterns with regard to clinical information. We also consider the option to customize the first mentioned visualization through settings (Annotation 3) to facilitate the radiomics analysis. For this we provide a dropdown containing patient-related categories such as histology or age and also a checkbox for highlighting potential clusters found in the radiomics data.

The wireframe shown in Figure 4.4 illustrates the user interface for the radiomics cluster analysis. Again, we plan two visualizations, this time for representing the cluster distribution (Annotation 4) and for displaying the radiomics feature distribution for one cluster (Annotation 5). Through these visual representations, we allow the observation of which patients exhibit similar radiomics feature values as well as the investigation of differences in feature values based on the segmentation. To further strengthen the cluster analysis, we provide information about the features which contributed to the cluster formation the most (Annotation 6). For this, we resort to the SHapley Additive exPlanations (SHAP) method [LL17].

Figure 4.5 depicts the wireframe of the UI component which specifically fulfills **Aim B**, which concerns probabilistic tumor segmentation and its effect on the radiomics analysis.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

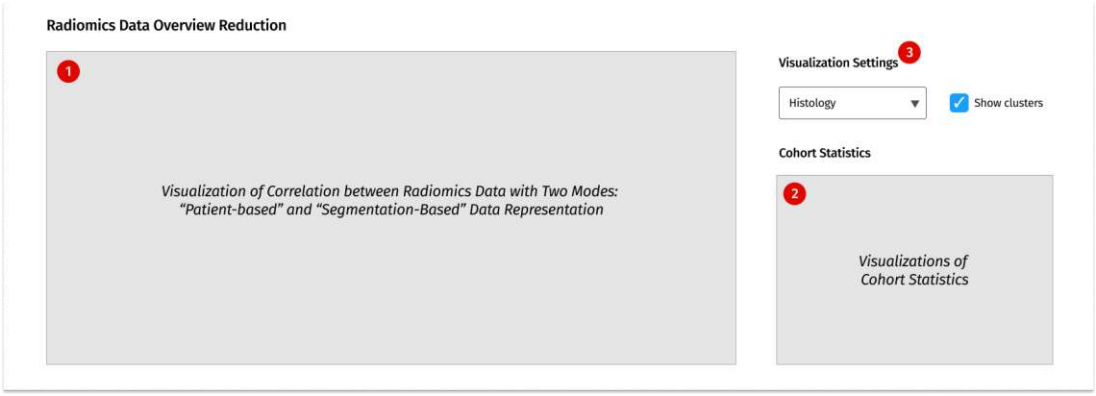


Figure 4.3: Wireframe of the UI component comprising the visualization of correlations within radiomics data (1), the respective settings (3), and the visualization of cohort statistics (2).

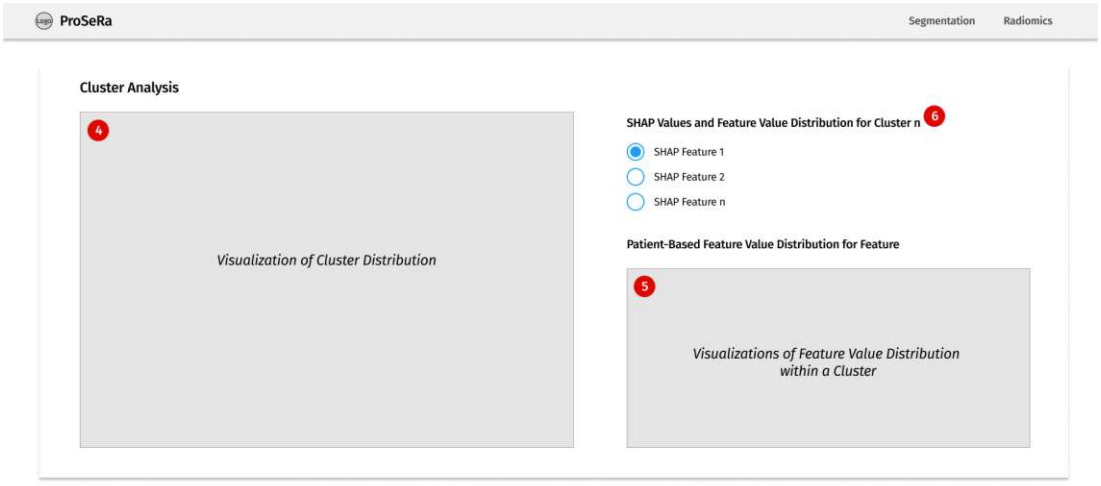


Figure 4.4: Wireframe of the UI component for the cluster analysis consisting of a visualization of the cluster distribution (4), the SHAP features (6), and a visualization of the distribution of feature values (5).

For this component, we plan one visualization which represents information about the segmentations based on segmentation probability (Annotation 7), which again is based on the interobserver variability present in our data set. In order to investigate the effect of delineation probability, we furthermore provide the option to select the desired probability threshold via a dropdown and a button to calculate the radiomics data based on the selection (Annotation 8).

The designed user interface enables user interactivity for the assessment of radiomics analysis of probabilistic tumor segmentations, as defined in **Aim C**. We consider cus-

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

tomizability options for the visualization of correlations between radiomics data through settings (Annotation 3 in Figure 4.3). This also applies to the list of SHAP features for each cluster (Annotation 6 in Figure 4.4), which empowers the users to select the feature of which they want to analyze the value distributions (Annotation 5 in Figure 4.4). Most importantly, we enable the selection of segmentation probability through a dropbox and the calculation of the radiomics analysis for the selected threshold, which is triggered by pressing a button (Annotation 8 in Figure 4.5).

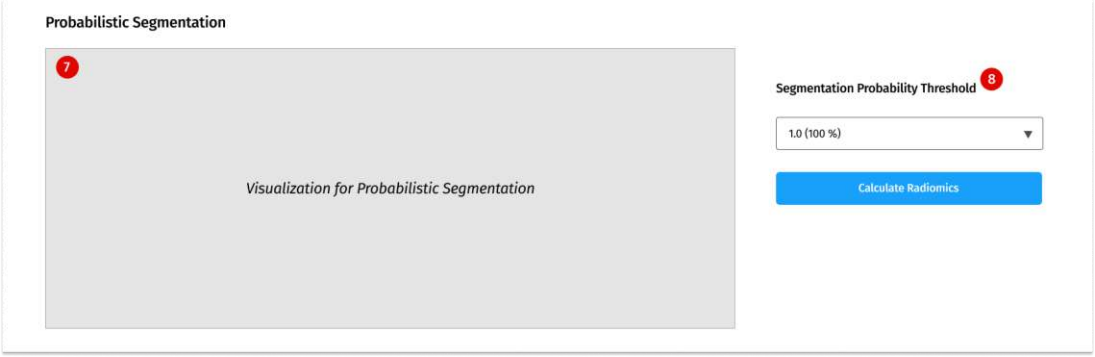


Figure 4.5: Wireframe of UI component for the analysis of probabilistic tumor segmentation (7) and the calculation of radiomics analysis based on probability threshold (8).

4.2.2 Visualization Design and Interaction

We strive for the visualizations to present valuable insights into the potential correlation of radiomics and clinical data and to show the effect of different tumor segmentations (**Aim A** and **B**) on the radiomics analysis. We anticipate the that the tumor assessments to have a considerable impact on the radiomics analysis. Thus, we update the radiomics information presented on the visualizations based on a chosen segmentation probability threshold. By means of user interaction, we trigger the visual interfaces to be updated (**Aim C**). In total, we plan five visualizations which are all linked with each other.

Design for Aim A: Visualization of radiomics with regard to clinical data

Aim A refers to the visualization of radiomics data with respect to clinical data. We aim to reach this goal by providing four visual representations focusing on different aspects of data. First, we show the *correlations within radiomics data* between patients to show potential patterns in our results. We strengthen the analysis of finding similarities and connections in our cohort by providing clinical information about the patient. Second, we display *statistics of the patient cohort* in another visualization. By this, we convey information about the patients regarding categories such as histology, age, sex, and overall tumor stage so that the users get a deeper insight into our entire cohort data. Third, we present the *distribution of clusters* as part of our cluster analysis. In our case, a cluster is generally a group of patients which are similar in terms of their calculated

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

radiomic feature values. We choose to visualize the cluster distribution to mainly provide an overview of clusters which among other benefits, facilitates the comparison between similar and also dissimilar patients. Last, we design a visualization which presents the *distribution of radiomics feature values* for each patient. We opt for this to show the different calculated values resulting from different segmentations related to each patient.

In general, there are several alternatives to show the correlation of data in terms of visualizations. Depending on the dimensionality, we can for example use correlation matrices and scatterplots to show pairwise correlations or PCP to show multidimensional correlations, similar to Yu et al. [YJY+17] or Moerth et al. [MWLH+20]. Considering that we want to visualize radiomics data that contain hundreds of features or dimensions, one may consider visualizations only suited for high dimensionality such as PCP. However, we are not restricted to these per se. As a solution, we can also perform dimensionality reduction on the radiomics data and obtain a lower dimensionality, usually limited to two or three dimensions. In this way, we can visualize high-dimensional data using visual representations such as scatterplots, which are usually used for presenting pairwise correlations. Additionally, this approach also reduces the complexity of the analysis as well, as such visualizations appear simpler and clearer.

We decide to represent the correlations within radiomics data by means of scatterplots, using two "modes" of representation: One mode shows the radiomics data within a *patient context* referring to all patients in the cohort, whereas the other mode focuses on the *segmentation context* for an individual patient. This allows a hierarchical representation of our cohort data, without compromising simplicity. Thus, we reduce complexity of the multidimensional radiomics data by performing dimensionality reduction, which we discuss in more detail in Subsection 4.3.2, prior to visualizing with a scatterplot. This results in a clear overview of data in form of data points. Each data point in a scatterplot either represents a patient (patient-based mode) among other patients or a segmentation (segmentation-based mode) among other segmentations for one particular patient. By clicking on a data point in the patient-based mode, the user can view the segmentation-based radiomics data of the particular patient. In order to get back to the patient-based mode, the user can click on the "Back to Overview" button. We illustrate our idea for the scatterplot in Figure 4.6.

In order to provide additional information about the patient cohort as well as the segmentations, we resort to the usage of colors and tooltips. We color-encode the data points in the scatterplot based on clinical data, particularly sex, age group, histology, and overall stage for the patient-based context, while for the segmentation-based context we color-encode based on the radiation oncologist and segmentation type. Thereby, the data points only correspond to one category at the same time, which can be easily identified through its assigned color. We also use different color schemes for each of these categories and make them consistent with the visualization for the patient cohort statistics which will be described later. This is done for aesthetic reasons, but also to prevent confusion between different categorizations. Further, we provide additional textual information to the data points by means of tooltips. When hovering over a data point, we show

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

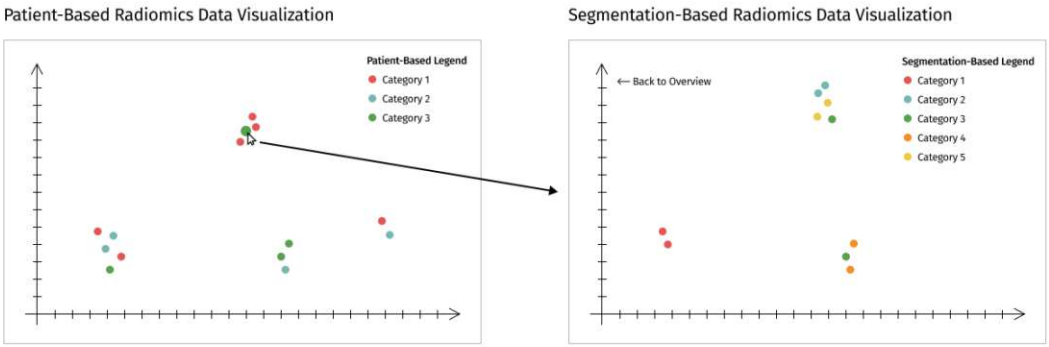


Figure 4.6: Idea for the patient-based and segmentation-based mode of the radiomics data visualization showing correlations within our data. The modes can be switched by clicking on one data point (patient) in the patient-based mode to view the segmentation-based data for the selected patient, or by clicking on the "Back to Overview" button in the segmentation-based mode.

the exact coordinates (x and y value) of the data point. Depending on the mode of the scatterplot, we display the information which is conveyed by means of color as text. With the tooltips, we want to avoid information overload on the scatterplot.

Considering the statistical analysis of the patient cohort, we provide visual representations combined with textual information. Examples of statistical visualization graphs include bar charts and doughnut charts, which are a variation of pie charts ([HR07]). One advantage of bar charts over doughnut charts may be the simple comparison between data based on the bar height compared to angles of a circle ([Dan14]). However, doughnut charts may be more suitable for representing data as part of a whole ([Dan14]).

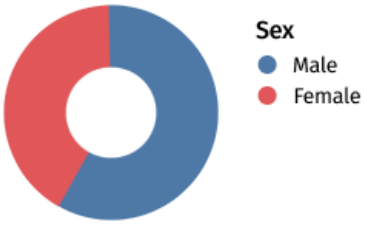
We ultimately choose to use **doughnut charts** to emphasize each portion belonging to a whole, or in this context, the patient cohort. Each doughnut chart represents the distribution of data for each clinical category, resulting in a total of four visualizations for sex, age group, histology, and overall stage. For each doughnut chart, we choose to add the list of categories including its percentage of the whole in a descending manner. This should specifically facilitate the comparison between parts of the data, while it also supports chart readability.

As previously mentioned, we incorporate colors into the doughnut charts to represent the clinical information about our patient cohort. In order to ensure consistency, we apply the same color schemes used in the scatterplot, which represents correlations within radiomics data. For the color encodings of all categories, we use the color schemes provided by D3[7], which include all scales from the commonly-used tool ColorBrewer[8]. Based on which clinical information we present, we use different color scale types, as

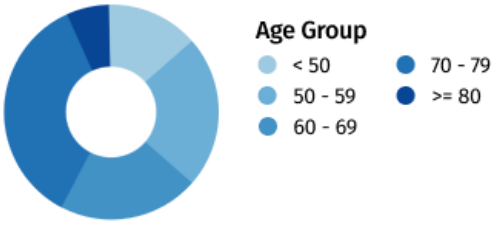
[7]D3 Color Scheme: `https://observablehq.com/@d3/color-schemes`
[8]ColorBrewer: `https://colorbrewer2.org`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

illustrated in Figure 4.7. Due to the categories "(biological) sex", "histology" and "overall stage" representing nominal data, we utilize qualitative or categorical color schemes which comprise clearly distinguishable colors (Figure 4.7a). These are appropriate for this type of data as they do not imply an implicit order, which is particularly important for nominal data [RTB96, SSM11]. Concerning the category "age group", we resort to sequential color scales. Due to these implying order, these are suitable for ordinal data which are usually related to numerical values such as age [SSM11]. For this, we utilize a sequential blue color scheme, with the lightest blue shade designating the youngest and the darkest blue shade designating the oldest age group (Figure 4.7b).



(a) Categorical or qualitative color scale for nominal data (e.g., sex).



(b) Sequential color scale for ordinal data (e.g., age).

Figure 4.7: Design of doughnut charts and respective legends in two different color scales showing different types of data.

Considering that we anticipate to find similarities or trends between patients or segmentations within the radiomics data, we provide visualizations for supporting the cluster analysis. First, we visually represent the structure of the cluster distribution. Alternatives for visualizing such data include heatmaps [YJY+17] and dendrograms [MEH+22]. Second, we provide insight about the distribution of the calculated radiomics feature value for each segmentation belonging to one patient within a selected cluster. For this case, boxplots or violin plots may be used [SWRT14]. Due to dendrograms and boxplots being more common and more straightforward visualization types in terms of simplicity compared to the other suggested alternatives, we use these for our cluster analysis.

We choose a **dendrogram** for visualizing the cluster distribution to provide a simple overview of which patients belong to which cluster (Figure 4.8). In our case, we consider a cluster a group of similar patients regarding their calculated radiomics feature values. In the visualization, one node represents one patient cluster, whereas the area size of the node is also matched to the size of the cluster, meaning its number of patients. Regarding the tree-like structure of the dendrogram, this means that the root is equivalent to the total patient cohort, whereas the leaves, which are the last and smallest nodes in the structure, mirror one patient. In a broader sense, we may regard one patient as a "cluster" of segmentation, considering our data consists of multiple segmentations for each patient resulting in multiple feature values.

In addition to the cluster distribution, we provide information about the validity of clusters.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

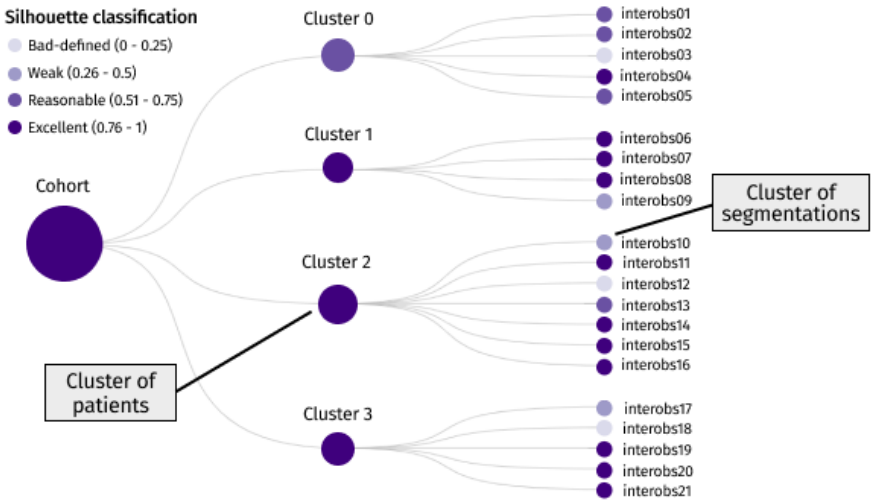


Figure 4.8: Design of dendrogram including the color encodings for representing the silhouette classifications and appropriate sizing based on the number of patients in a cluster. The last nodes in the structure (leaves) signify a "cluster" of segmentations of one patient, whereas the other nodes signify a group of patients (actual cluster).

We do this to enable the comparison between clusters regarding their internal structure. Similar to the approach by Raidou et al. [RvdHD+15] presented in Subsection 3.2.3, we use the metric *average silhouette coefficient* ($s = \frac{BSS-WSS}{max(BSS,WSS)}$), which describes how cohesive one cluster is (WSS) compared to the other clusters (BSS) [TSK05]. In our context, the average silhouette coefficient represents the cohesion of one patient in terms of the segmentations compared to the other patients within the same patient cluster. This means that we calculate s for one patient, who we consider "cluster" of ten segmentations, with the exception of one patient (`interobs19`), which consists of five segmentations. An illustration of our concept is represented in (Figure 4.9). For each patient cluster (consisting of several patients), we calculate the mean of the silhouette values for each patient within the cluster with $\overline{s_C} = \frac{1}{n_C} \sum_{p \in C} s_p$, where C signifies a cluster of patients, n_C the number of patient in a cluster, p a patient in the cluster C, and s_p the silhouette coefficient for each patient within the cluster.

Due to the silhouette coefficient (s) values being numerical, we apply a sequential color scale, similarly to the scheme used for the age group category in the scatterplot. As mentioned in Subsection 3.2.3, the values for s range between 0 to 1 [TSK05]. We categorize these values into four groups as presented by Raidou et al. [RvdHD+15], which are: (1) $0 - 0.25$: bad-defined, (2) $0.26 - 0.5$: weak, (3) $0.51 - 0.75$: reasonable, and (4) $0.76 - 1$: excellent. Based on these definitions, we color-encode each node with a sequential purple color scale. As shown in Figure 4.8, bad-defined cluster nodes are filled with the lightest purple shade and the excellent cluster nodes with the darkest

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

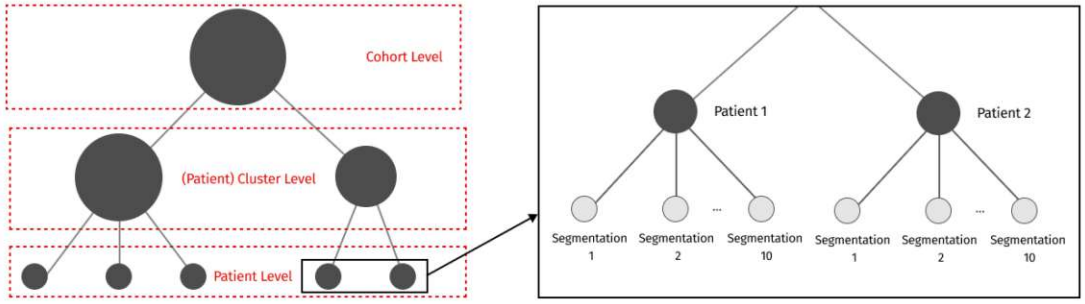


Figure 4.9: Concept of dendrogram structure (left) with a detailed illustration of one patient cluster consisting of multiple patients, whereas each patient is considered a "cluster" of segmentations (right).

purple shade. We choose purple in order to avoid the usage of the same color schemes for different visualizations focusing different topics. Besides the color-encodings, we again utilize tooltips to convey information about the size and silhouette coefficient of one cluster.

Since the scatterplot only presents the radiomics data after dimensionality reduction, we employ a **boxplot** which shows the distribution of values for each of the radiomic features. We opt for a **boxplot** containing several groups and showing individual data points (Figure 4.10). In this context, this means that the visualization shows the value distribution of one feature for all segmentations (i.e., point) of each patient (i.e., group). As this visual interface is part of the cluster analysis, we want the boxplot to enable the comparison of feature value distribution between patients within *one* cluster. Regarding the dimension, the x-axis corresponds to the patients in one cluster, whereas the y-axis is mapped to the values of one radiomics feature.

In order to further strengthen the cluster analysis, we also calculate the five most important SHapley Additive exPlanations (SHAP) features [LL17] for each cluster. SHAP is a unified framework which supports the interpretations of predictions of a model by determining the importance value for each feature [LL17]. SHAP values therefore express how much each feature contributes to the formation of a model, which, in our case, is a cluster. We decide to solely provide boxplots for these so that we avoid information overflow and the user can focus on the features which are most prominent within a cluster.

Design for Aim B: Investigation of the effect of probabilistic tumor segmentation on the radiomics analysis

The goal of investigating probabilistic segmentation and the effect on the radiomics analysis is stated with **Aim B**. Specifically for the probabilistic segmentation task, we provide one visualization representing the total share of pixels with which the segmentations coincide. In other words, we show the agreement among oncologist observers

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

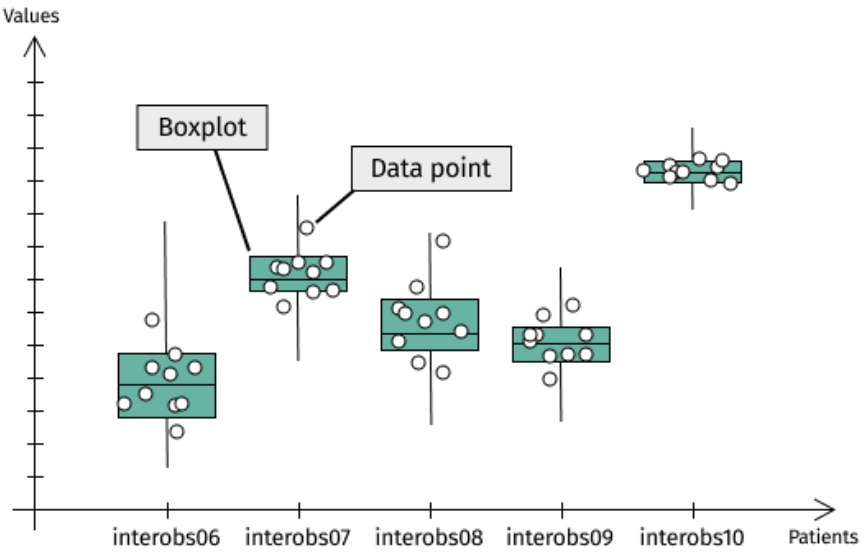


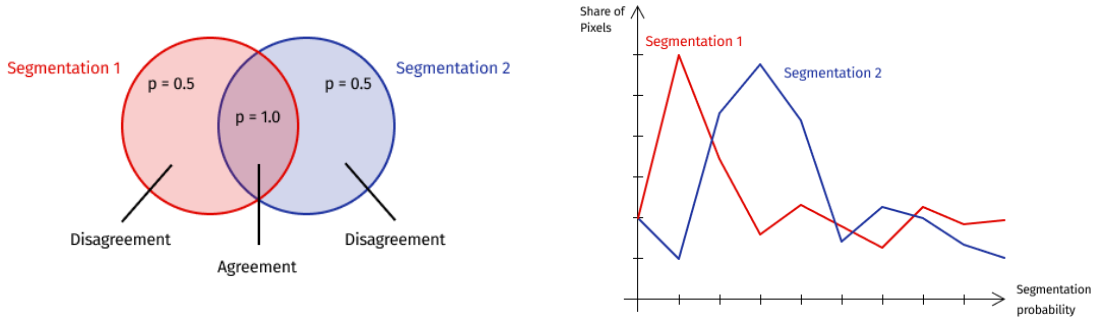
Figure 4.10: Design of boxplot with each group representing the value distribution of one selected feature for one patient within a cluster. Data points (white dots) are added for each group to show the exact value for each segmentation.

regarding which pixels represent part of a tumor, which can be an indication of a robust segmentation. Alternatives for the visualization of this purpose include line or area charts ([SSM+16, DSD+21]). When the data includes several groups to be represented each, line charts may be more suitable as the overlap of colored areas in area charts may not convey meaningful information.

Due to the larger number of patients, we use a **line chart** in which one line represents the share of pixels over the tumor segmentation probability of one patient. We visually illustrate our idea in Figure 4.11, which shows the concept of segmentation accuracy (Figure 4.11a) and of the line chart (Figure 4.11b). Whereas the x-axis represents the tumor probability and ranges from 0.1 (= 10%) to 1.0 (= 100%), the y-axis correlates with the share of pixels (in absolute numbers). As we generally have ten segmentations for each patient, we decide to look at the probability in 0.1 steps. This results in a straightforward approach regarding the calculation of the share of pixels, since one segmentation represents a 0.1 probability, given that we have 10 segmentations in total.

Considering the impact on the radiomics analysis, we provide the option to interactively switch the segmentation probability threshold one wants to observe and explore the effect of this change to the analysis in real-time. Based on this setting, the components for the radiomics analysis described in Section 4.2.2 are updated. As this feature relates to the goal of providing user interaction stated with **Aim C**, it will be described in the following section in more detail.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) Venn diagramm for the segmentation probability concept.

(b) Concept for line chart representing the share of pixels over the segmentation probability.

Figure 4.11: Concept for the visualization of probabilistic tumor segmentation illustrated in an example with two segmentations.

Design for Aim C: User Interaction

Aim C focuses on providing user interaction specifically for assessing the impact of probabilistic tumor segmentation on the radiomics analysis. In order to reach this goal, we incorporate interactive features within our visualizations. The main interaction techniques in *ProSera* are *filtering*, which enables the observation of a data subset based on filter criteria, [Spe07], *brushing and linking*, which allows the examination of subsets within a data in another linked visualization [Kei02], and *focus-plus-context*, which enables the view of one object in detail while still providing contextual information [CMS99]. These empower the customization of our visual interfaces, resulting in the opportunity of thoroughly analyze our results based on the desired focus aspects. In Figure 4.12, we illustrate the used interaction techniques which we describe in detail in this section.

Due to our results being comprised of large amounts of information, we provide *filtering* to enable the observation of our data based on selected criteria. We do this in our scatterplot, for which we allow the selection of patient category (biological sex, age group, histology, overall stage) and segmentation category (oncologist, segmentation type). The selection triggers changing the conveyed information based on the chosen category, hence also modifying the color schemes of the data points.

We incorporate the technique of *brushing and linking* for most of our visual interfaces. In the line chart representing probabilistic segmentation, we enable the option of brushing through the probability thresholds to highlight the data of interest. Further, the selection of one segmentation accuracy threshold of interest triggers updating the data represented in all visualizations for the radiomics analysis based on the threshold (linking). We also include the functionality in our scatterplot representing correlations of radiomics data. As we have already mentioned, this representation comprises two modes, the patient-based and the segmentation-based mode. When clicking on one data point representing one patient in the patient-based mode, the scatterplot switches into the segmentation-based mode for the selected patient. We also implement brushing and linking in the visual

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

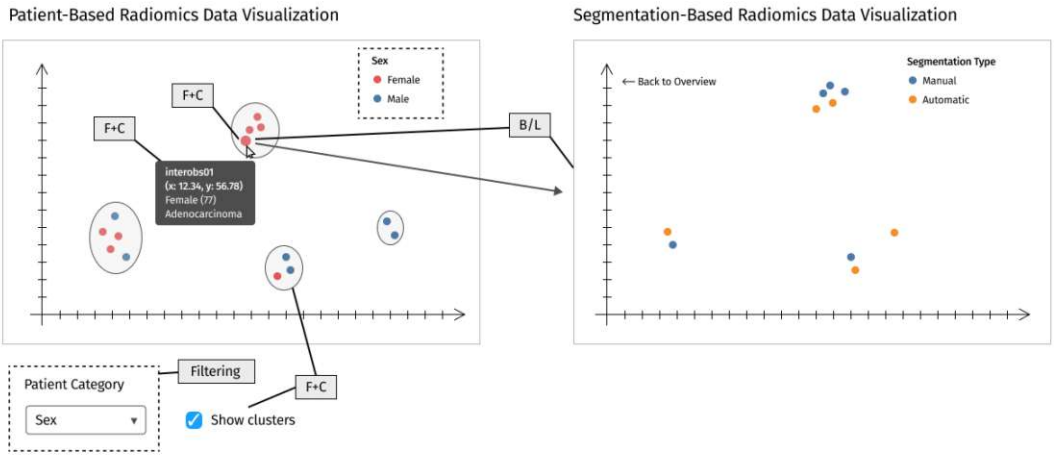


Figure 4.12: Interaction techniques used in our scatterplot: *filtering* of patient categories, *focus-plus-context (F+C)* by means of focused data points, tooltips and cluster markings, and *brushing and linking (B/L)* through clicking on one data point (patient) to switch to the segmentation-based view of the selected patient.

interfaces for the cluster analysis. By clicking on the nodes representing one patient cluster in the cluster distribution dendrogram, the cluster context for boxplot should be updated as well. This also applies to the list of most significant SHAP features, which differ for each cluster. This list in turn corresponds to the boxplot as well, as the selection of one feature triggers updating the boxplot regarding the distribution of values for the specific feature.

We employ the *focus-plus-context* technique in order to provide detailed information of one object of interest in an overview. We do this specifically by displaying tooltips when hovering over data points in the scatterplot, nodes in the dendrogram, and data points and boxplots in the radiomic feature distribution visualization. For example, the tooltip for a data point in the patient-based scatterplot shows information about the patient, the respective clinical data and the x and y-coordinates of the data point. Furthermore, we also indicate the focus of one object by visually highlighting it from the other objects in the same context. We apply this in our scatterplot and the boxplot, where we increase the size of a data point when the user hovers over it. Similarly, we highlight the nodes representing one patient cluster by encircling them with a border. In addition to the increased size of focused elements, we also include the option to show or hide cluster information in the scatterplot.

4.3 Implementation

We ultimately provide an “offline” web application regarding the computation of radiomics data. In other words, we calculated all data for the radiomics analysis for each probabilistic

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

threshold in advance. The main reason for this decision was the fast retrievability of radiomics data, resulting in an instant information availability. This method allows a fast response from the web application which would have been much slower if the radiomics data were calculated on demand. In the following sections, we describe the implementation procedure for the tasks for data preprocessing followed by the user interface and visualizations.

4.3.1 Radiomics Feature Extraction

As core part of the radiomics analysis, we extracted the radiomic features from medical imaging. For this purpose, we used the open-source Python package `PyRadiomics` [vGFP+17b], particularly the "Feature Extractor" class. The package supports several radiomics feature classes, including first order, shape, GLCM, and GLSZM.

Before we started with the extraction, we needed to prepare the segmentation and medical imaging data especially regarding compatibility with the extractor. In this case, we transformed DICOM files into nearly raw raster data (NRRD) files using the open source software 3D Slicer[9] as well as the Python packages `Pydicom`[10] and `Pydicom-Seg`[11]. We used the first mentioned tool to convert all 2D CT slices of each patient into 3D volumetric data but also to gain an initial insight into the input data. By means of the `Pydicom` and `Pydicom-Seg` packages we then read the DICOM segmentation files for each patient and created an NRRD file for each annotated segmentation.

We extracted the radiomic features from the generated NRRD files using the PyRadiomics package. As seen in Listing Listing 4.1, we initialized the radiomics feature extractor with the setting `geometryTolerance`. This was necessary due to a slight image/mask geometry mismatch potentially caused by small deviations in space in the used data set. Since we used a very small value for the increase of geometry tolerance, we expect no negative consequences on the validity of the radiomics analysis. We saved the extracted radiomics values of all segmentations of all patients into a single CSV file. We also added patient clinical information for each row for the designing the visualization later. In general, the CSV file consists of thirty-two dimensions (patient, sex, age, histology, tumor location, T-stage, N-Stage, M-Stage, overall stage, segmentation, 22 different diagnostic information such as original image size) and 107 domains equivalent to 107 radiomic features. In total, the file consists of 205 rows (excluding the header), with each row designated to one segmentation for one patient. It should be noted that the 22 dimensions for diagnostic information were provided by the feature extractor by default. These values were not relevant and therefore not taken into account for the radiomics analysis.

[9] 3D Slicer: `https://www.slicer.org`
[10] Pydicom: `https://pydicom.github.io`
[11] Pydicom-Seg: `https://razorx89.github.io/pydicom-seg`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Listing 4.1: Initialization of Radiomics Feature Extractor

```
from pyradiomics import featureextractor

settings = {'geometryTolerance': 0.0001}
extractor = featureextractor
    .RadiomicsFeatureExtractor(**settings)
```

4.3.2 Dimensionality Reduction

With regard to reducing dimensionality within our generated data, we also consider each patient as an entity besides looking at each segmentation separately. Thus, we employed data wrangling on our CSV file containing the extracted radiomic features. Data wrangling denotes the process of transforming data to prepare them for analysis [FGL+16]. Regarding our data specifically, we "merged" each row belonging to one patient into one row. In order to avoid duplicate column variables, we only appended the radiomics feature values to each patient row and appended the segmentation annotation to the name of the radiomics feature. As a result, we ended up with 21 rows for the "patient-based" CSV file, whereas one row is designated for one patient.

In order to find the most appropriate reduction method for the input data, we investigated three dimensionality reduction techniques, which are: (1) Principal Component Analysis (PCA), (2) t-distributed Stochastic Neighbor Embedding (t-SNE), and (3) Uniform Manifold Approximation and Projection (UMAP). Whilst PCA is a technique which preserves the global structure of data regarding distance, t-SNE preserves the local distances [PHPQACPQ20]. UMAP manages to balance between preserving global and local distances by including consistent and stable structures [PHPQACPQ20]. As the range of values regarding all radiomics feature values altogether is quite large, we have *normalized* these values prior to dimensionality reduction.

We utilized the `Scikit-learn` software machine learning library [PVG+11] and UMAP-learn [MHM18] for implementing the dimensionality reduction. We have used the "StandardScaler" class from the `sklearn.preprocessing` package to normalize the radiomics feature values. For performing PCA and t-SNE, we have used the "PCA" class from the `sklearn.decomposition` package and the "TSNE" class from the `sklearn.manifold` package, respectively. As the name says, we have used UMAP-learn for performing the UMAP. Aside from the reduction technique itself, the appropriate parameters for these algorithms are also of great importance. As we have anticipated t-SNE and UMAP to fit the nature of our data the most, we have also compared different values for the critical parameters *perplexity* (t-SNE) and *neighbors* (UMAP). The values were adjusted based on quality of the results. All results from our investigation can be seen in Figure 4.13. It can be seen that UMAP with the parameter neighbor set to two produced the most meaningful data. Compared to the other results, a desired formation of visual clusters is visible. As already mentioned, we planned to reduce the dimensions

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

for both patient-based as well as segmentation-based data. Again, the results were stored in CSV files.

The final step of the dimensionality reduction task consisted of applying a cluster algorithm on our data to retrieve meaningful groupings of patients. Alternatives for clustering include k-Means [HE03] and density-based spatial clustering of applications with noise (DBSCAN) [EKSX96]. Whereas the first is based on partition regarding different groups in data, the latter is based on density regarding space. We have ultimately decided to use k-Means due to it being partition-based, as we anticipate our data to form clusters based on specific traits such as sex, histology or age. We again resorted to `Scikit-learn` for clustering data, particularly the "KMeans" class from the `sklearn.cluster` package. Finally, we added the resulting cluster information to the dimensionality reduction data.

4.3.3 Cluster Analysis

For the cluster analysis, we evaluated the dimensionality reduction results including the generated clustering data by means of simple metrics. In our case, these help to facilitate the comparison between clusters and patients within a cluster. Common metrics include the Hopkin's statistic [HS54], which checks the clusterability of projections regardless of cluster labels, and silhouette coefficient or score [Rou87], which validates quality of cluster formation based on known clusters.

We use the silhouette coefficient metric since we have already defined data clusters using UMAP dimensionality reduction [MHM18] followed by k-Means clustering [HE03]. The simplicity and straightforward characteristic of this metric also contributed to our decision. We manually calculated the silhouette coefficients in Python. For the calculation of the average silhouette coefficient, we have regarded one patient as a "cluster" of segmentations [TSK05]:

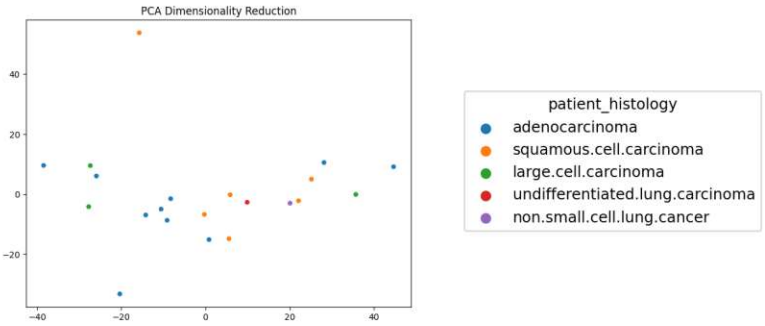
$$s_p = \frac{BSS - WSS}{max(BSS, WSS)} \tag{4.1}$$

where p represents a patient, BSS the separation, which shows the distinction of the patient within a cluster of patients, and WSS the cohesion of the patient regarding its segmentations. We discussed these metrics in detail in Subsection 3.2.3. We then calculated the mean of all patient silhouettes within each cluster:

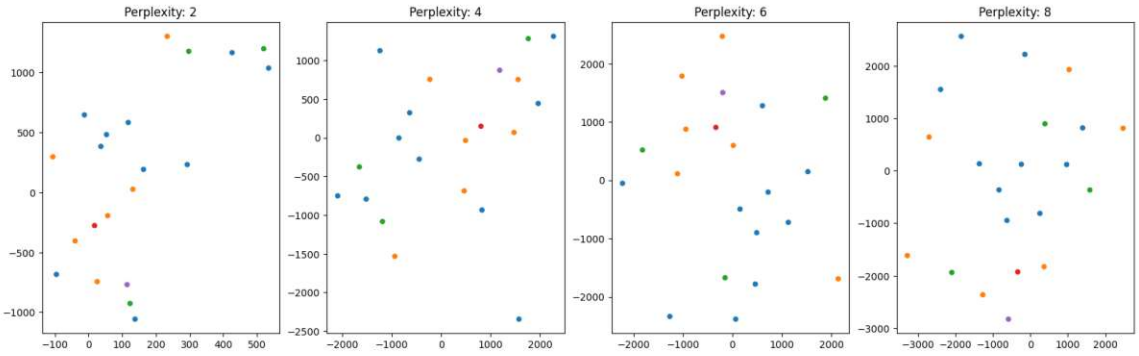
$$\overline{s_C} = \frac{1}{n_C} \sum_{p \in C} s_p \tag{4.2}$$

where C reflects a cluster of patients, n_C the number of patients in the selected cluster, p the patient in a cluster C and s_p the silhouette coefficient for a patient. It should be noted that we use a linear metric for on the results generated by the non-linear approach UMAP, however, we did again employ the linear k-Means algorithm on these results.

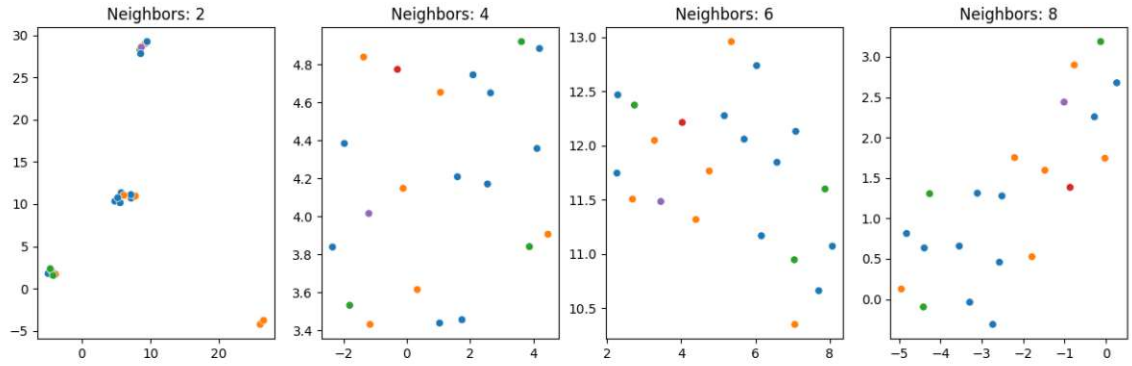
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



(a) Results for PCA.



(b) Results for t-SNE with perplexity 2, 4, 5, and 6.



(c) Results for UMAP using neighbors 2, 4, 6, and 8.

Figure 4.13: Results of different dimensionality reduction approaches on patient-based radiomic features using different parameterization. The legend seen in (a) applies to all plots.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

As we wanted to further define our the clusters context of radiomic features, we chose to determine the features which contributed most to the formation of each cluster. In this way, we aim to empower the observers to focus on a much smaller amount of radiomic features. For this purpose, we calculated the SHAP [LL17] values for every cluster, which gives us information about the most significant radiomic features. We did the calculations in Python by using the `SHAP` package[12]. The SHAP value determination signifies the last task regarding calculations for the cluster analysis.

4.3.4 Probabilistic Tumor Segmentations

In order to obtain tumor segmentation data based on probability, we compared the binary masks of all segmentations for every patient. In other words, we superimposed all segmentations for each patient based and calculated the tumor probability of each pixel. We did this by counting the number of agreements of one pixel belonging to one tumor (meaning, where one pixel in the binary mask equals to 1) and divided by the number of segmentations conducted for the patient. In our case, we have a total of 10 segmentations for each patient, with the exception of one patient (`interobs19`), for whom five segmentations are available. Based on this approach, we generated one NRRD file for each patient containing the probability of one pixel being a tumor based on the provided segmentation data.

By means of the results from the masks comparison, we generated probability density function (PDF) for each patient so that the share of pixels for each probability threshold is evident. Therefore, we have omitted the background within the images by only considering the pixel values which are not 0. We then calculated the PDF using the formula:

$$P_t = \frac{n_t}{N_p} \tag{4.3}$$

where P_t represents the probability function for a user-selected threshold t, n_t represents the number of pixels with probability threshold t, and N_p represents the total number of pixels which potentially designate a tumor for patient p. We define this as the union of pixels in all tumor segmentations for one patient:

$$N_p = s_1 \cup s_2 \cup ... \cup s_n \tag{4.4}$$

where s_n signifies the (number of pixels for) n-th tumor segmentation for one patient. As we have generally have ten segmentations for each patient, the PDF for one patient is represented by a 10×1 matrix. For example, the PDF for `interobs05` is defined as:

$t:0.1$	$t:0.2$	$t:0.3$	$t:0.4$	$t:0.5$	$t:0.6$	$t:0.7$	$t:0.8$	$t:0.9$	$t:1.0$
$\lceil 0.17$	0.10	0.04	0.03	0.03	0.02	0.03	0.03	0.05	$0.49 \rceil$

[12]SHAP package: `https://shap.readthedocs.io`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

For instance in this example, there is an 100% agreement on 49% of the segmented pixels belonging to a tumor.

As we have gained information about the segmentation probabilities, we then focused on providing segmentation data for the radiomics analysis for each probability threshold. First, we determined which "step" size between each treshold we wanted. Since we generally have ten segmentations for each patient, we consequently decided on ten thresholds, with the step size being 0.1 (or 10%) for the range from 0 to 1. Second, we generated masks or "new" segmentations for each probability threshold. We did this by again comparing all pixels from the binary masks but only including the ones which apply to the specific threshold value. In total, we have generated 205×10 segmentations, resulting in 2050 NRRD segmentation files. We have then used these files as input for performing the radiomics analysis for each probability threshold and saving the results accordingly.

There were few obstacles concerning the extraction of radiomics features from of the probabilistic segmentation data. As some segmentation files only contained a very small number of segmented pixels, the radiomics extraction was not always feasible. In some cases, some radiomic features could not be extracted, whereas in other cases, no radiomic features could be extracted at all. We resolved these issues by simply removing all calculations for one segmentation if it the radiomics feature extraction was incomplete. We have also done this due to these entries not being comparable anyway due to lacking data. Thereby, it should be noted that not all radiomic features extraction files include the calculations of all segmentations. This is caused by some segmentations simply not having segmented pixels to which certain probability thresholds apply.

4.3.5 User Interface and Visualizations

The implementation of the visual analytics tool involved designing visualizations and also preprocessing data, specifically radiomics and segmentation data. This means that we also dealt with the extraction of radiomic features, the dimensionality reduction of radiomics data, the analysis of clusters, and the calculation of probabilistic tumor segmentations. The technology stack for fulfilling these tasks stack includes D3.js, which was used for the visualization task, Angular, which was used for the frontend or user interface, and Python, which was used for the data processing task. We resorted to PyCharm and Visual Studio Code as our integrated development environment (IDE) for the calculation and visualization tasks, respectively.

As already mentioned, we used the framework Angular together with D3.js to implement the user interface as well as the visualizations. For the UI elements specifically, we have utilized the UI component library Angular Material[13]. Figure 4.14, Figure 4.15, and Figure 4.16 show screenshots of *ProSeRa*, which consists of the following:

[13] Angular Material: `https://material.angular.io`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

1. **Probabilistic segmentation component** (Figure 4.14), including the toolbar or header of the web application and the probabilistic segmentation settings.

2. **Dimensionality reduction component** (Figure 4.15), including banner showing the selected probabilistic threshold, the scatterplot and respective settings.

3. **Cluster analysis component** (Figure 4.16), including the dendrogram for the cluster distribution, the list of calculated SHAP values for the selected cluster, and the boxplot for the feature value distributions.

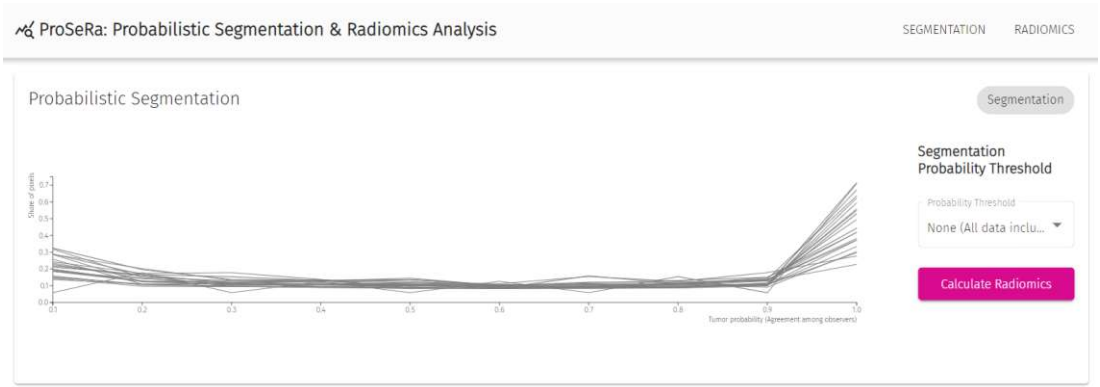


Figure 4.14: Screenshot of application toolbar and probabilistic segmentation component including line chart and probability threshold setting.

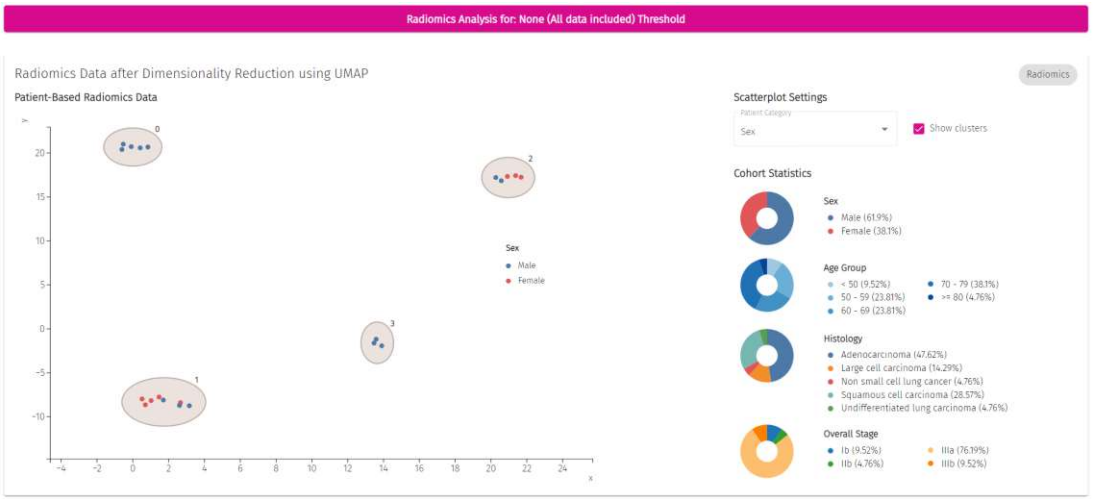


Figure 4.15: Screenshot of probabilistic threshold information and dimensionality reduction component including scatterplot (patient-based data), doughnut charts (cohort statistics), and scatterplot settings.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

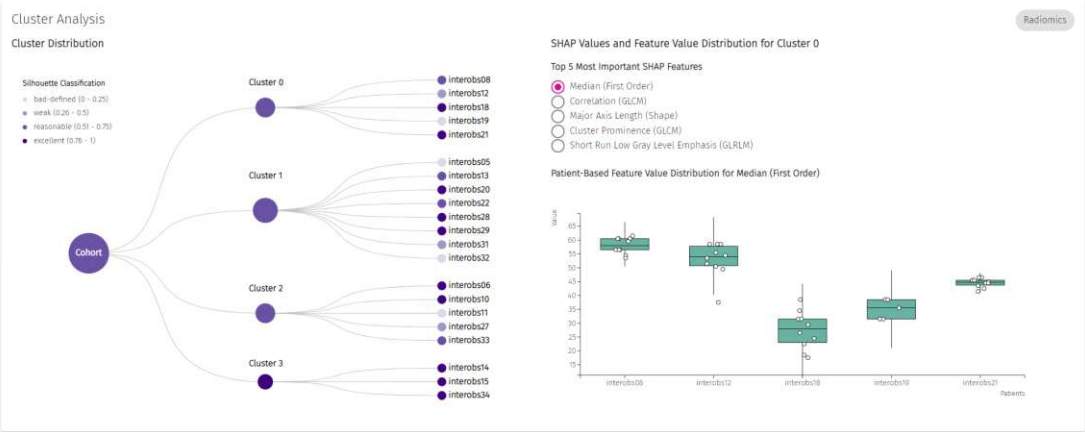


Figure 4.16: Screenshot of cluster analysis component including dendrogram (cluster distribution) and boxplot (feature distribution).

All visualizations for the radiomics analysis are triggered by changing the probability threshold in the probabilistic segmentation component. This can be done in two ways: By simply selecting the desired threshold in the dropdown or by clicking on the line chart, as seen in Figure 4.17. When hovering over the line chart with the mouse, the rectangle showing the considered probability threshold changes according to the x-point of the mouse. By clicking on the "Calculate Radiomics" button, the radiomics analysis component are updated based on the selected threshold.

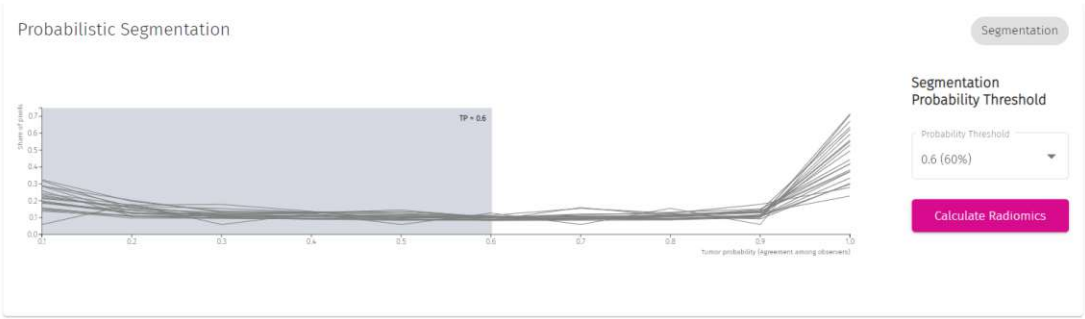


Figure 4.17: Brushing and linking interaction of the line chart to select the probability threshold.

As we already mentioned in Subsection 4.2.2, the scatterplot in the dimensionality reduction component provides two modes: The patient-based mode, where one point represents one patient, and the segmentation-based mode, where one point represents one segmentation, as seen in represented in Figure 4.18. We additionally implemented tooltips when hovering over the points of interest and an overview of patient information, including displaying one CT scan of the patient. In order to leave this view, we further

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

provided a “Back to Overview” button to head back to the patient-based scatterplot mode.

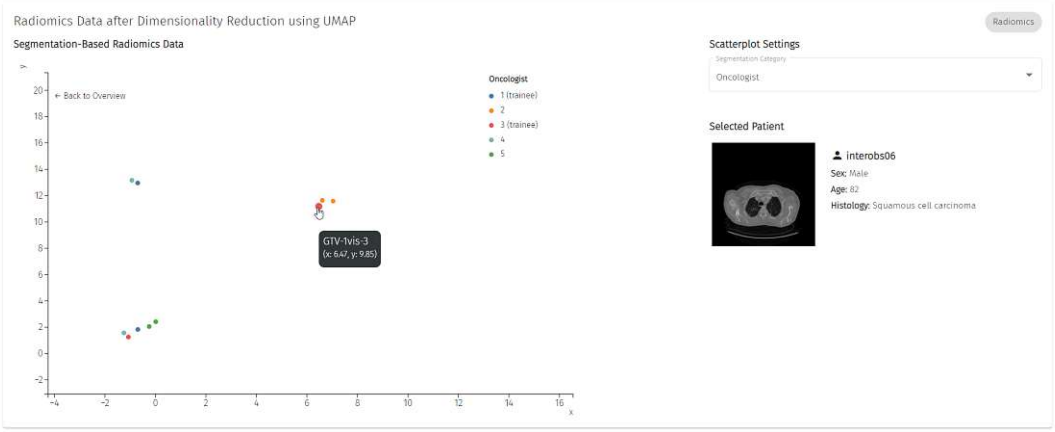


Figure 4.18: Segmentation-based scatterplot view for `interobs06`, including tooltips (brushing and linking) and display of patient information (focus-plus-context).

Considering the cluster analysis component, we also implemented interaction features within the visualizations. By clicking on a cluster circle in the dendrogram, the boxplot including the list of most important SHAP features are updated. The boxplot is also updated according to the selected SHAP feature for which the distribution of values should be displayed. Similar to the scatterplot, we implemented a tooltip for both dendrogram and boxplot when hovering over a circle or a box, respectively. The tooltip provides detailed information about the hovered element which may not be visible by the naked eye such as the statistical values of a boxplot.

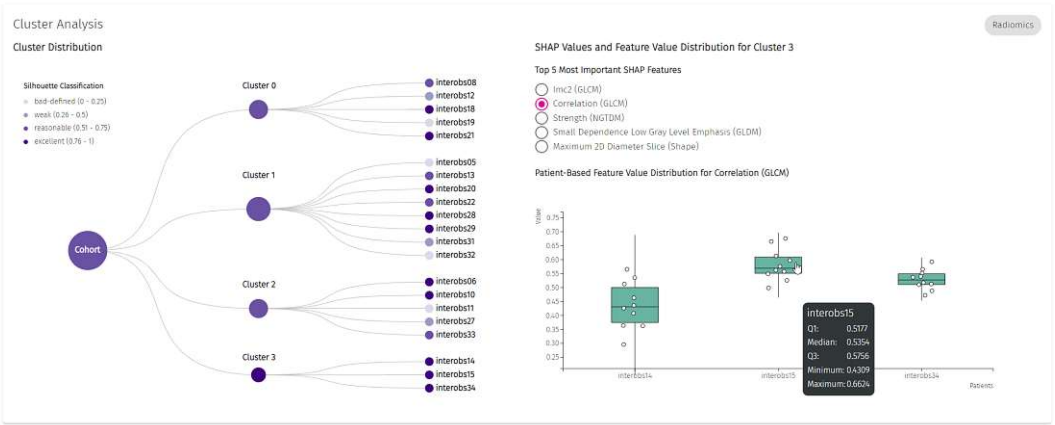


Figure 4.19: Cluster analysis visualizations after changing cluster through brushing and linking and SHAP feature and including a tooltip in the boxplot (focus-plus-context).

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 5

Results and Discussion

In this chapter, we present our results using five usage scenarios which we defined based on significant and unexpected findings. We evaluate the outcomes using evaluation methods appropriate for answering our research question. Subsequently, we discuss our approach followed by stating the limitations of our work.

5.1 Results

Prior to presenting our results, we want to state and briefly describe alternatives for evaluation methods and which of these we plan to apply. The literature proposes several approaches to evaluate information visualizations such as the nested model for visualization design and validation [Mun09], Visual Data Analysis and Reasoning (VDAR)[LBI+12], and Qualitative Result Inspection (QRI). [IIC+13]. Munzner [Mun09] introduces a nested model which supports the selection of appropriate evaluation methods. For each of the four nested layers, the model determines the threats to validity and suggests evaluation approaches to overcome these. Lam et al. [LBI+12] propose VDAR, which determines the support of a visualization tool on the quality of data analysis including knowledge gain in a specific domain. It assesses the contribution to supporting data exploration, knowledge discovery, hypothesis generation, and decision-making processes by methods such as case studies. Isenberg et al. [IIC+13] present QRI assessments which concentrate on evaluating and discussing visualization results in a qualitative manner. In contrast to VDAR it focuses on validating the image quality, visual encodings, and system behavior or interaction (walkthrough).

As we target the evaluation of our results based on how well the interactive visualizations in *ProSeRa* support the radiomics analysis, we chose VDAR [LBI+12] and QRI [IIC+13] as our evaluation scenarios. In the context of our work, this means that we assess the visual analytics tool based on (1) its quality in supporting the investigation of probabilistic

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

segmentation on the radiomics analysis and (2) the quality of its visualizations. We evaluate our results by means of usage scenarios on interesting and remarkable findings.

5.1.1 Usage Scenario 1: Correlation between Tumor Segmentation Probability and Cluster Formation

As we anticipated an influence of tumor segmentation probabilities on clustering, we particularly investigated the size of clusters and a potential categorization of clusters based on clinical data. We therefore compared how the visualizations of the correlations (scatterplot) are impacted by different tumor segmentations. We came to the findings that (1) a 100% probability relates to the highest cluster formation, and (2) the clinical categorization of clusters not being completely blatant.

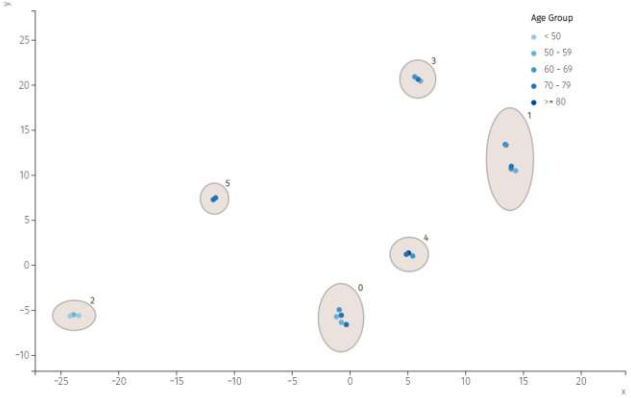
Generally, the segmentation probability was found to closely correlate to the formation of clusters as well as the respective cluster sizes. As seen in Table 5.1, the cluster size does not increase continuously with an increasing or decreasing segmentation probability, meaning that there is not direct or inverse proportionality. However, the highest amount of clusters was indeed formed with a 100% segmentation probability, with the smallest cluster consisting of two, and the biggest cluster consisting of four patients. The second most number of clusters was found at an 80% probability, with six clusters having at least two and at most five patients. Apart from these thresholds, the number of clusters range from four to five, with the smallest minimum cluster size being two and the highest maximum cluster size being ten.

Regarding the characterization of clusters based on clinical data, we could not find any clear patterns concerning histology or age group. In Figure 5.1, we provide screenshots of patient-based radiomics visualization with age group and histology as classification category for 80% (Figure 5.1a and Figure 5.1b) and 100% (Figure 5.1c and Figure 5.1d). It can be seen that most clusters cannot be linked to one particular classification for age group or histology. In case for the age group, cluster 0 in Figure 5.1a for example includes patients from the age 50 up to 79. There are minor exceptions where some clusters only comprise patients of one age group, such as cluster 1 and 5 to which patients from the age group 70 – 79 belong. For the most part, however, clusters cannot clearly be linked to one age group. This is also the case for histology, were the distribution among clusters is even more dispersed. Nevertheless, we did expect this particular behavior for this category. This is due to some of the histology classifications being subtypes of others, such as squamous cell carcinomas, adenocarcinomas, and large cell carcinomas being subtypes of NSCLC. Also, the fact that adenocarcinomas can be found in each cluster can be explained through the fact that adenocarcinomas make up 40% of the NSCLC [TBN+15].

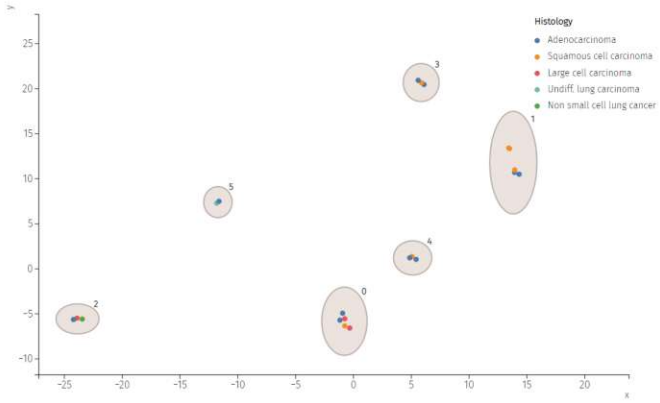
Despite not finding any patterns for the classification of clusters according to age and histology, we did discover some clusters which could be assigned one classification regarding biological sex and overall tumor stage. As seen in Figure 5.2, four out of six clusters (Figure 5.2a), and four out of seven clusters (Figure 5.2c) can be classified as

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

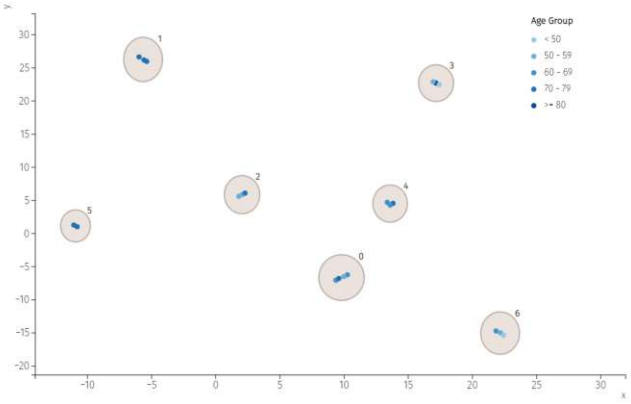
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



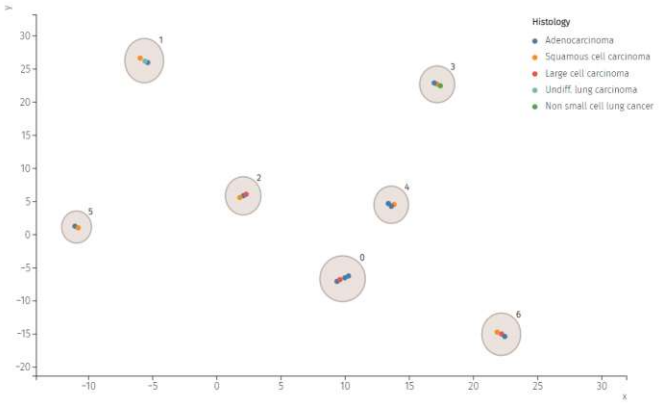
(a) Age group at 80% probability



(b) Histology at 80% probability



(c) Age group at 100% probability



(d) Histology at 100% probability

Figure 5.1: Cluster classifications in patient-based radiomics data based on “age group” and “histology”.

Table 5.1: Overview of number of clusters and minimum and maximum cluster size regarding tumor segmentation probability.

Segmentation Probability	Number of Clusters	Minimum Cluster Size	Maximum Cluster Size
-	4	3	8
0.1 (10%)	5	3	6
0.2 (20%)	4	4	7
0.3 (30%)	5	3	5
0.4 (40%)	5	2	8
0.5 (50%)	4	2	7
0.6 (60%)	5	2	10
0.7 (70%)	4	3	8
0.8 (80%)	6	2	5
0.9 (90%)	5	2	7
1.0 (100%)	7	2	4

entirely male or female at 80% and 100% tumor probability, respectively. Similarly, three out of six clusters (Figure 5.2b) and four out of seven clusters (Figure 5.2d) comprise patients diagnosed with stage III (IIIa and IIIb) cancer. As a result, we can deduce that the tumor phenotypes for lung cancer may look similar or produce similar radiomics feature values among patients with the same sex or overall tumor stage.

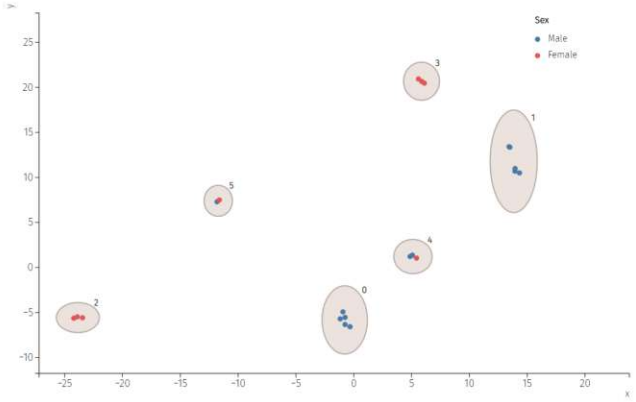
5.1.2 Usage Scenario 2: Correlation between Tumor Segmentation Probability and Cluster Silhouettes

In addition to exploring the formation of clusters, we also wanted to investigate the cluster silhouette based on tumor segmentation probability. In doing so, we have compared the visualization for showing the cluster distribution (dendrogram) for different probability thresholds. We specifically focused on the calculated silhouette values representing how well a cluster is defined.

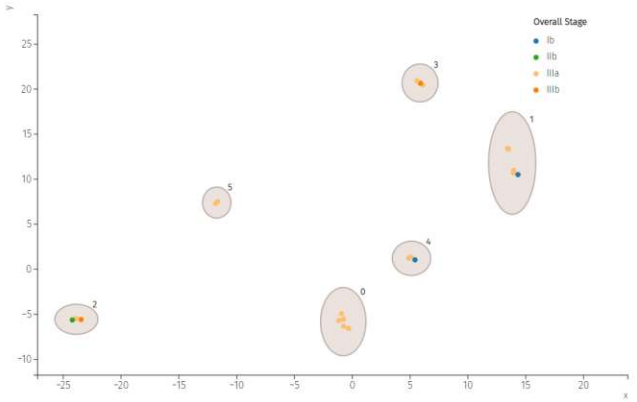
Similar to Usage Scenario 1, the cluster definition is neither directly nor inversely proportional to the segmentation probability threshold. Figure 5.3 compares the cluster distribution including the patient silhouettes at different probabilities. It shows that the silhouette of one patient cluster is not necessarily dependent on the probability threshold, with 40% segmentation probability producing generally excellently-defined clusters (Figure 5.3a), 60% producing worse defined clusters in comparison (Figure 5.3b),

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

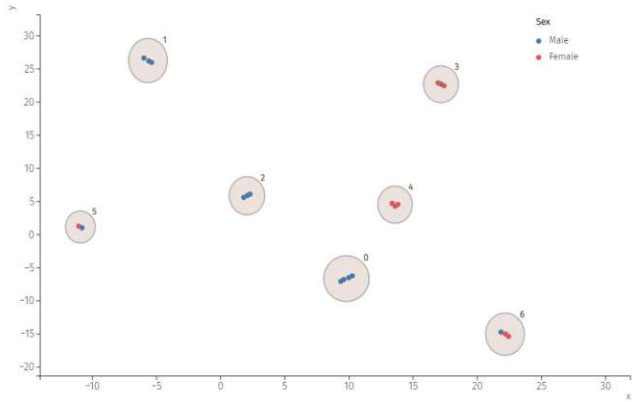
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



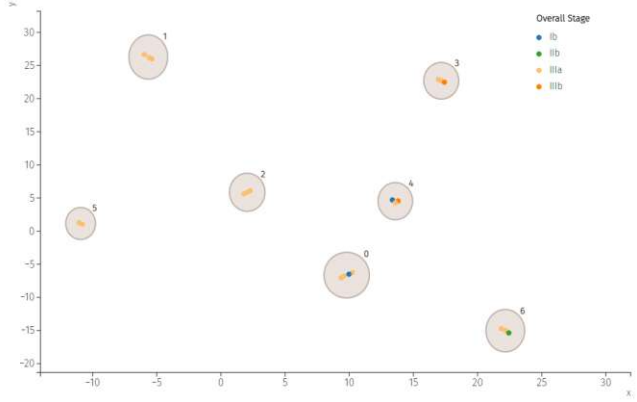
(a) Sex at 80% probability



(b) Overall stage at 80% probability



(c) Sex at 100% probability



(d) Overall stage at 100% probability

Figure 5.2: Cluster classifications in patient-based radiomics data based on "(biological) sex" and "overall (tumor) stage".

and 80% producing better defined clusters again (Figure 5.3c). Of course, 100% segmentation probability resulted in solely excellently defined clusters (Figure 5.3d). Since this threshold includes no discrepancies within the segmented tumors, the silhouettes for each patient cluster equal to 1.0.

Despite not detecting any patterns regarding proportionality, we did observe occasional abnormalities within clusters regarding the silhouette classifications. Occasionally, there are occurrences of bad-defined patients among otherwise mostly well- or excellently-defined patients within the same cluster (Figure 5.4). For example, this is the case for cluster 0 and 2 at a 50% segmentation probability (Figure 5.4a) or for cluster 1 and 3 at a 70% probability (Figure 5.4b). We explain this as follows: (1) The concerning patients deviate from other patients within the clusters based on radiomics feature values, or (2) there was a high interobserver variability in the tumor segmentation for the concerning patient. The last reason may apply to patient interobs15, who produced bad-defined silhouette values at 50% threshold ($s_{15} = 0.25$), 70% threshold ($s_{15} = 0.03$), and 80% threshold ($s_{15} = 0.15$).

Furthermore, we also discovered a profound influence of probabilistic segmentation onto the silhouette classifications. In Figure 5.5, we have compared the cluster distribution generated from all segmentation data (Figure 5.5a) to the distribution generated with the consideration of a 60% threshold, which produced the worst values compared to other thresholds (Figure 5.5b). When looking at the silhouette values calculated from all input data, meaning that no threshold is selected, we can observe three reasonably-defined (75%) and one excellently-defined cluster (25%). This is significantly worse compared to the silhouette classifications computed for any probability threshold, which exhibited at least 40% of clusters being excellent-defined (60% probability threshold). More details will be provided in Usage Scenario 3.

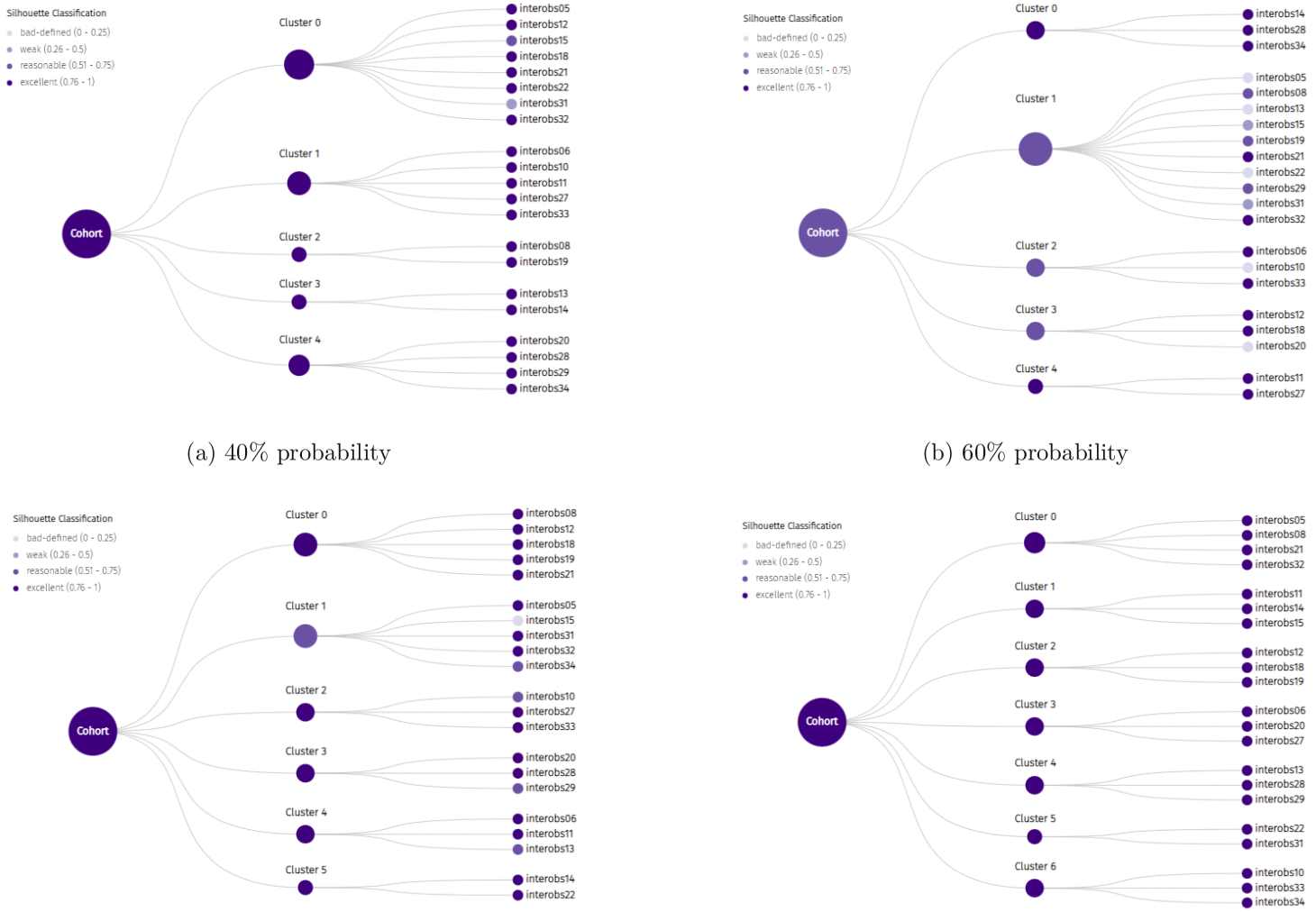
5.1.3 Usage Scenario 3: Radiomics Analysis When Including All Tumor Segmentations

With this usage scenarios we wanted to provide a thorough analysis of radiomics data calculated from all segmentation data. We investigated all visualization components so that we can gain insight on the original input data without taking the probability threshold into consideration. In this way, we can explore the degree of interobserver variability within patients' segmentations as well as its effect on the radiomics feature extraction.

The radiomics data after dimensionality reduction resulted in four clusters separating 21 patients (Figure 5.6). When looking at the inter-cluster level, we can see that the clusters are very distant from one another. This means that these can be considered being quite distinct in terms of radiomics feature values. At the same time, this also signifies that patients could be grouped based on the radiomics data extracted from their tumor segmentations, hinting at similar or similar-looking phenotypes. When observing the data within the clusters itself (intra- cluster level), however, we could not find a clear

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) 40% probability

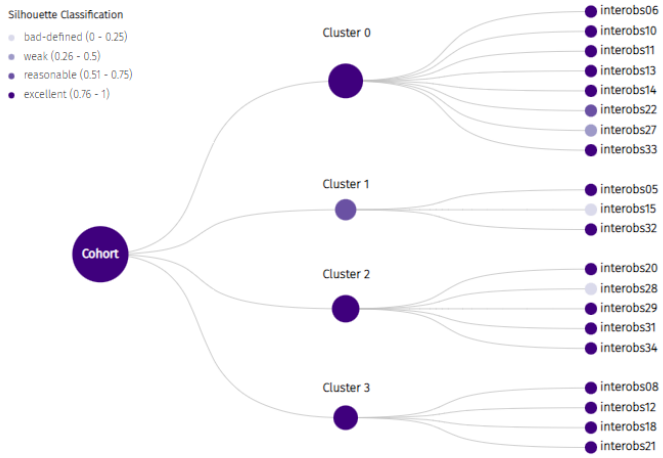
(b) 60% probability

(c) 80% probability

(d) 100% probability

Figure 5.3: Comparison of cluster silhouette classifications at different probabilities .

(a) 50% probability



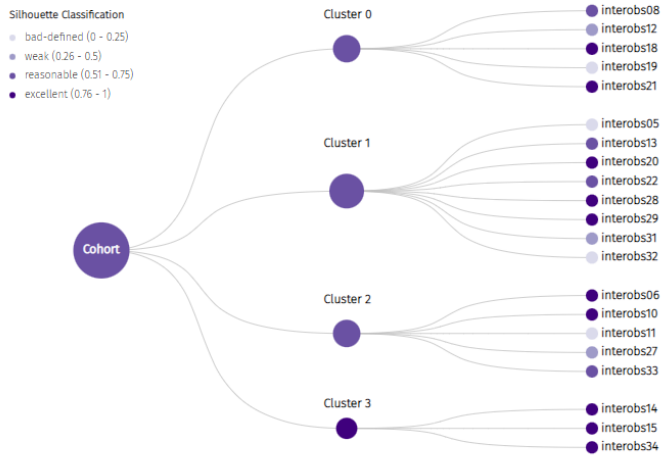
(b) 70% probability

Figure 5.4: Comparison of cluster silhouette classifications at different probabilities.

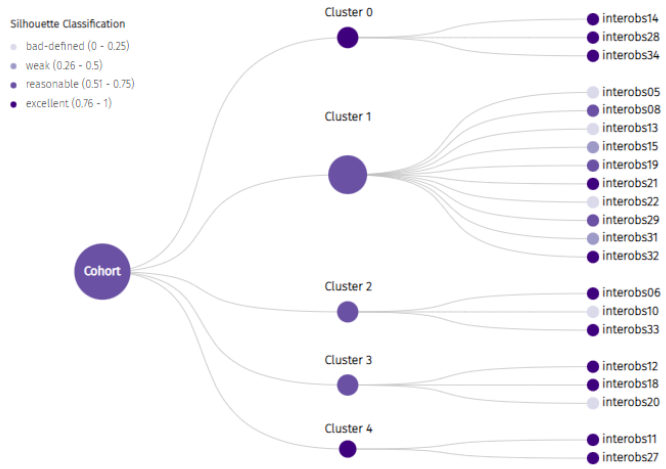
coherence among the patients regarding histology (Figure 5.6a) or age (Figure 5.6b). Nevertheless, we could categorize some clusters based on biological sex. As seen in Figure 5.6c and Figure 5.6d, cluster 0 and 3 solely consists of male patients with stage IIIa lung cancer. This indicates that NSCLC patients of the same sex diagnosed with the same cancer stage may show similarities regarding their tumors and tumor phenotypes.

As we also wanted to examine on the separation and coherence (i.e., silhouette) of clusters, we observed both the just described scatterplot (Figure 5.6) and the dendogram showing the silhouette classifications (Figure 5.7). Overall, the cluster sizes range from three to

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) No threshold (All data included)



(b) 60% probability threshold

Figure 5.5: Comparison of cluster silhouette classifications regarding the consideration of probability threshold.

eight patients per cluster, with cluster 1 being the largest. The mean of silhouettes for each cluster are defined as "reasonable"($0.55 <= \overline{s_C} <= 0.63$) except for cluster 3, which is categorized as "excellent" ($\overline{s_3} = 0.85$). This particular cluster is the smallest one with three patients, who are all male, between 69 – 77 old and were diagnosed with either an undifferentiated lung carcinoma or a squamous cell carcinoma. Compared to the other clusters, cluster 3 can be considered the most coherent in terms of clinical patient data. In contrast to this, cluster 2 can be considered the most incoherent cluster 2 with

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

(a) Histology (b) Age group

(c) Sex (d) Overall stage

Figure 5.6: Radiomics data after dimensionality reduction based on different patient categories.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

$\overline{s_2} = 0.55$. This is not just evident from the silhouette values shown in the dendrogram but also from the clinical patient data in the scatterplot.

Figure 5.7: Cluster distribution including silhouette classification for all input data.

We further observed the feature value distributions of SHAP features by means of the boxplot (Figure 5.8). In doing so, we specifically examined the feature distributions for "Dependence Variance (Gray-Level Dependence Matrix (GLDM))" for the most incoherent cluster, cluster 2. As expected, there are conspicuous variances within each patient's segmentations (white data points). Especially patient `interobs27` shows a great variance with values between 14 and 25 and the white data points being separated the most compared to the ones for the other patients. In contrast, `interobs33` shows the least variances but has one noticeable outlier located above all other values. This means that the one segmentation, in this case the manual segmentation by oncologist 2 (as seen in the tooltip in Figure 5.8), significantly diverges from the other segmentations in terms of the Dependence Variance feature. Moreover, we also noticed a variance in the feature values between the patients as well, as each boxplot is plotted at different heights.

5.1.4 Usage Scenario 4: Consistency and Coherence of Cluster Formation Regarding Tumor Segmentation Probability

By means of usage scenarios 4 we aimed to investigate how different segmentation probabilities influence the stratification of our data. We specifically were interested in whether clusters stay together different throughout accuracy thresholds. In this way, we also aspired to detect "strong" clusters so that we can deduce which patients are very similar to one another regarding their radiomics feature values.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

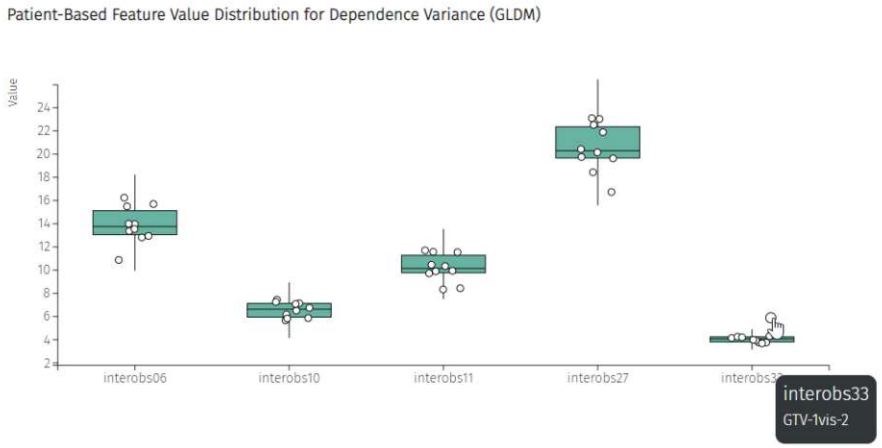


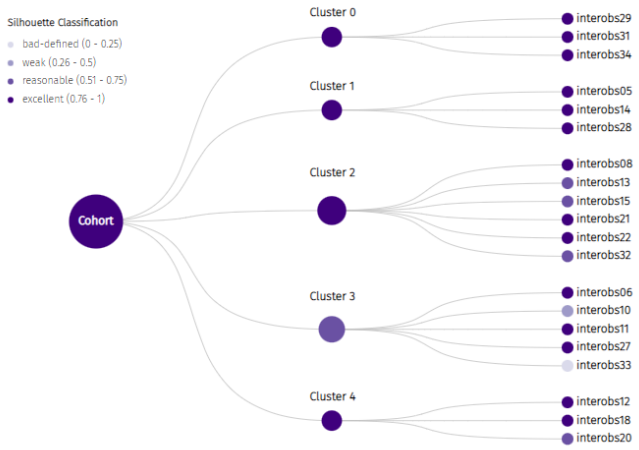
Figure 5.8: Distributions for "Dependence Variance (GLDM)" values for cluster 2, with a tooltip showing segmentation information about the outlier for `interobs33`.

Altogether, we found the segmentation probability having a tremendous impact on the stratification (clustering). Figure 5.9 shows this impact by comparing the cluster distribution for a 10% segmentation accuracy (Figure 5.9a) to a 90% segmentation accuracy (Figure 5.9b). Despite having the same number of clusters formed, the allocation of patients to the clusters is significantly different for some cases. For example, patients `interobs11` and interobs15 form one cluster (cluster 4) at 90% accuracy but assigned to different clusters (cluster 2 and 3) at 10% accuracy. Similar behavior is also reflected for cluster distributions for other probability thresholds.

However, there is one group of patients which particularly stood out to us regarding cohesion. `Interobs06`, `10`, `11`, `27`, and `33` (Cluster 3 in Figure 5.9a) always formed one cluster for the probability thresholds 10% – 50% as well as when there was no probability threshold set. At 60%, the cluster was split into two clusters, with `interobs11` and `interobs27` being divided from the other patients. Similar behavior is also shown for a 70% – 90% accuracy, with one or more completely different patients being occassionally added to the (sub-)clusters. At 100% segmentation accuracy, the initial cluster was divided into three clusters, with other patients being added to these. The patients which were consistently put into the same cluster were `interobs10` (Female, 47, NSCLC, IIIa) and `interobs30` (Male, 58, Adenocarcinoma, IIIa). We provided an overview in Figure 5.10, comparing the cluster distribution at 20% (Figure 5.10a), 50% (Figure 5.10b), 80% (Figure 5.10c), and 100% accuracy (Figure 5.10d).

Even though the cluster formation for `interobs06`, `10`, `11`, `27`, and `33` stayed coherent for the most part, the SHAP features for the cluster(s) with respect to different segmentation accuracy diverged significantly. For example, the clusters for the patients generated with 10% and 50% segmentation threshold exhibit completely distinct SHAP features, despite containing the exact same patients. There were some overlaps such as

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN



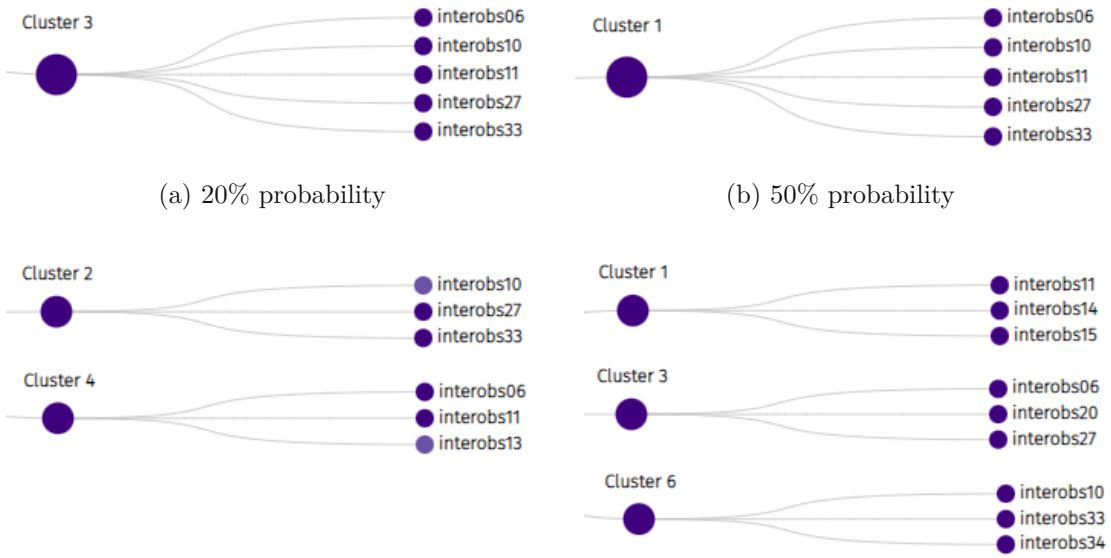
(a) 10% probability

(b) 90% probability

Figure 5.9: Cluster formation for different probability thresholds with same number of clusters but different cluster allocation.

the "Maximum 3D Diameter" feature being listed in the SHAP feature list for a 10% and 80% probability. However, these were only limited in number. The clusters 1, 3, and 6, which include `interobs06`, `10`, `11`, `27`, or `33` at 100% tumor probability, do not have any (top five most important) SHAP features in common. Nevertheless, we could detect that most of these features belong to the category GLRLM and GLDM. We provided an overview of the most important SHAP features for each cluster which includes `interobs06`, `10`, `11`, `27`, and `33` at 100% segmentation probability in Table 5.2. We also illustrate the boxplots for the most prominent SHAP features for each cluster in

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) 20% probability

(b) 50% probability

(c) 80% probability

(d) 100% probability

Figure 5.10: Cluster distribution for `interobs06`, `10`, `11`, `27`, and `33` at different accuracy thresholds.

Figure 5.11, with each boxplot showing the feature value distributions for each patient within the same cluster (Figure 5.11a, Figure 5.11b, and Figure 5.11c).

5.1.5 Usage Scenario 5: Patients with Low versus High Sensitivity to Changes in the Probability Threshold

Besides observing the influence of tumor segmentation probability on clustering, we wanted to focus on how individual patients are affected by segmentation accuracy. For that, we particularly explored the silhouette values as well as the segmentation-based radiomics data after dimensionality reduction and compared the results for each accuracy threshold. Generally, we did not find any patient in our data whose silhouette was consistently bad- or weakly defined. However, we did observe some patients being barely affected (low-sensitivity), while others being highly affected (high-sensitivity) regarding segmentation accuracy. We will present and discuss these through by means of one example for each category in the following paragraphs.

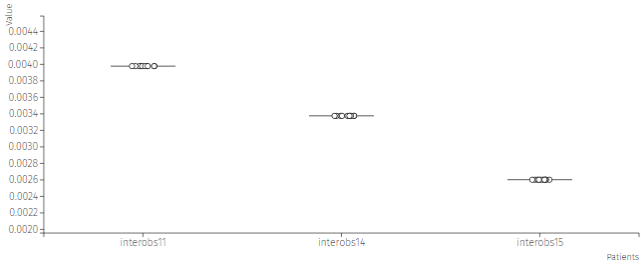
We start with introducing the patients with the lowest and highest sensitivity to provide insight into the clinical data as well as their range of silhouette values. `Interobs14` (Male, 77, Undifferentiated lung carcinoma, IIIa) appears to being affected by the segmentation accuracy the least. This patients exhibits an excellent silhouette for all probability thresholds, with values ranging from 0.92 (worst) to 1.0 (best), meaning

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

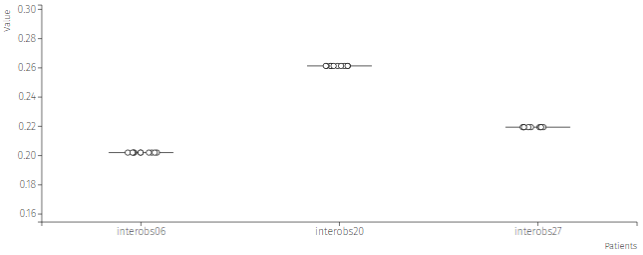
Table 5.2: SHAP features for clusters containing `interobs06`, `10`, `11`, `27`, and `33` (marked in bold) at 100% tumor segmentation probability.

Patients Cluster	Top Five Most Important SHAP Features
• **interobs11** • interobs14 • interobs15	• Small Dependence Low Gray Level Emphasis (GLDM) • Long Run Emphasis (GLRLM) • Long Run High Gray Level Emphasis (GLRLM) • Dependence Non Uniformity Normalized (GLDM) • Dependence Entropy (GLDM)
• **interobs06** • interobs20 • **interobs27**	• Dependence Non Uniformity Normalized (GLDM) • Long Run Low Gray Level Emphasis (GLRLM) • Small Area Low Gray Level Emphasis (GLSZM) • Dependence Variance (GLDM) • Surface Volume Ratio (Shape)
• **interobs10** • **interobs33** • interobs34	• Long Run Emphasis (GLRLM) • Run Length Non Uniformity (GLRLM) • Voxel Volume (Shape) • Gray Level Non Uniformity (GLRLM) • Long Run Low Gray Level Emphasis (GLRLM)

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) "Small Dependence Low Gray Level Emphasis (GLDM)" feature value distributions for cluster containing `interobs11`, `14`, and `15`.



(b) "Dependence Non Uniformity Normalized (GLDM)" feature value distributions for cluster containing `interobs06`, `20`, and `27`.

(c) "Long Run Emphasis (GLRLM)" feature value distributions for cluster containing `interobs10`, `33`, and `34`.

Figure 5.11: Boxplots showing the feature value distributions for the most prominent SHAP feature of clusters containing `interobs06`, `10`, `11`, `27`, or `33` at 100% probability (Table 5.2). Note here that the values for each segmentation are the same at 100% probability (no segmentation disagreement), hence each group in the boxplots being represented as a line.

that the values remain consistent. In contrast, `interobs31` (Male, 73, Squamous cell carcinoma, IIIa) shows a high variance in silhouette based on segmentation probability, with values ranging from 0.02 (worst) to 1.0 (best). The silhouette does not continuously increase with a higher segmentation accuracy but varies without any particular pattern. In Table 5.3 we have compared the silhouette values between these patients with regard to the segmentation probability. We illustrate the silhouette values for `interobs14`

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Table 5.3: Comparison of silhouette values between `interobs14` (low-sensitivity patient) and `interobs31` (high-sensitivity patient) with respect to segmentation probability.

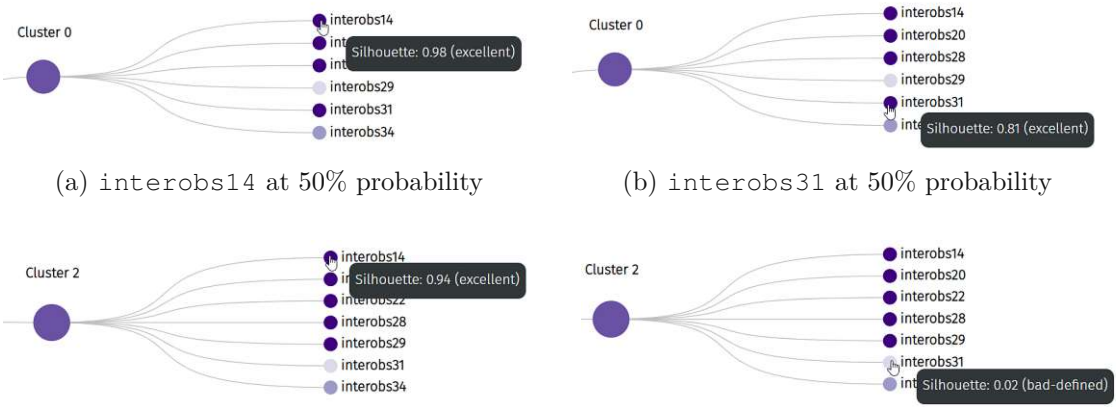
Segmentation Probability	Silhouette for `interobs14` (Low-sensitivity patient)	Silhouette for `interobs31` (High-sensitivity patient)
-	0.84	0.33
0.1 (10%)	1.0	0.86
0.2 (20%)	0.99	0.31
0.3 (30%)	0.99	1.00
0.4 (40%)	1.0	0.46
0.5 (50%)	0.98	0.81
0.6 (60%)	0.97	0.27
0.7 (70%)	0.99	0.97
0.8 (80%)	0.92	0.98
0.9 (90%)	0.94	0.02
1.0 (100%)	1.0	1.0

and `interobs31` at 50% and 90% shown in the cluster distribution dendrogram in Figure 5.12. Figure 5.12a and Figure 5.12b show the silhouette values for `interobs14` and `interobs31` at 50% segmentation probability, whereas Figure 5.12c and Figure 5.12d show the silhouette values for `interobs14` and `interobs31` at 90% segmentation probability.

We explain our findings with the following reasons: First, tumors of low-sensitivity patients may exhibit characteristics such as clear borders which are easier to assess compared to tumors of high-sensitivity patients. As a result, the correlated interobserver variability regarding the tumor segmentations will be accordingly lower or higher. We studied this aspect by observing the scatterplot visualizing segmentation-based radiomics data for each patient (see Figure 5.13). When comparing the results produced with all segmentation data (Figure 5.13a and Figure 5.13b), we can first and foremost see that both exhibit clusters containing similar segmentations. When glancing at the scale, however, we can see that the distance between the clusters differs. For `interobs31`, this distance is larger, meaning that groups of segmentations did not produce the same radiomics features, hinting at a higher interobserver variability. This is even more apparent at 60% probability threshold (Figure 5.13c and Figure 5.13d), where no clusters are apparent for the high-sensitivity patient in contrast to the low-sensitivity patient.

Second, the quality of medical imaging also may significantly influence the tumor segmen-

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



(a) `interobs14` at 50% probability

(b) `interobs31` at 50% probability

(c) `interobs14` at 90% probability

(d) `interobs31` at 90% probability

Figure 5.12: Silhouette values for `interobs14` (low sensitivity) and `31` (high sensitivity) shown in tooltips `11`, `27`, and `33` at different accuracy thresholds.

tation process. Factors such as low contrast aggravate the assessment of tumors which also impacts the variability among observers. In Figure 5.14, we have compared one CT scan each of the `interobs14`, which is the low-sensitivity patient (Figure 5.14a), and `interobs31`, which is the high-sensitivity patient (Figure 5.14b). The CT scans clearly show a significant difference in contrast, with the CT scan of `interobs31` exhibiting a much lower contrast compared to the CT scan of `interobs14`.

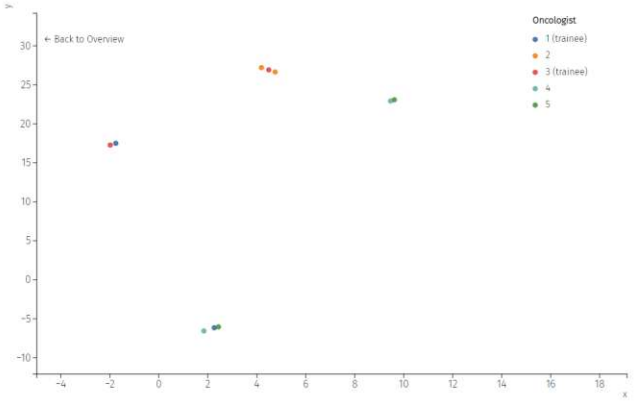
5.2 Discussion

The visual analytics tool strongly supported us in investigating the effect of tumor segmentations on the radiomics analysis. It empowered us to observe the calculated radiomics data for different segmentation probability thresholds by simply switching the respective settings. This enabled us to thoroughly analyze the result of each accuracy threshold and also to compare the results of different thresholds. By means of a dataset containing tumor segmentations conducted by actual oncologists, we could investigate a case which reflects interobserver variability in a real-world setting.

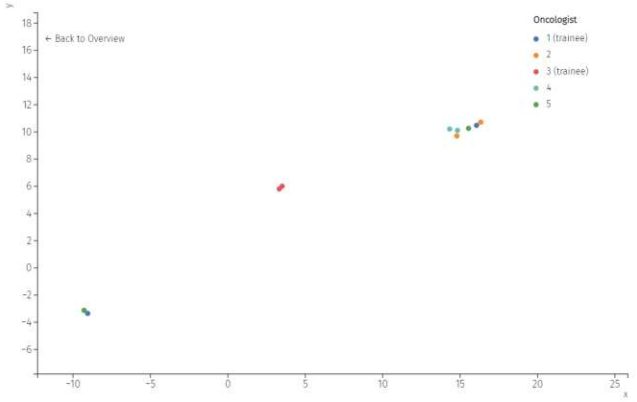
Even though radiomics data are known for their complexity, the visual interface enables an analysis of these in a rather clear and simple way. The preprocessing of radiomics data largely contributes to this aspect since we principally focused on extracting meaningful data but also on simplifying information. By this, we refer to tasks such as reducing the dimensionality of radiomics data after extraction or calculating the silhouette values as part of the cluster analysis. Besides the preprocessing process, the visualizations also support a facilitated radiomics analysis by means of appropriate visualization types for conveying the intended information. We specifically also strived for reducing information overload by using appropriately designed encodings and simple guidance methods.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

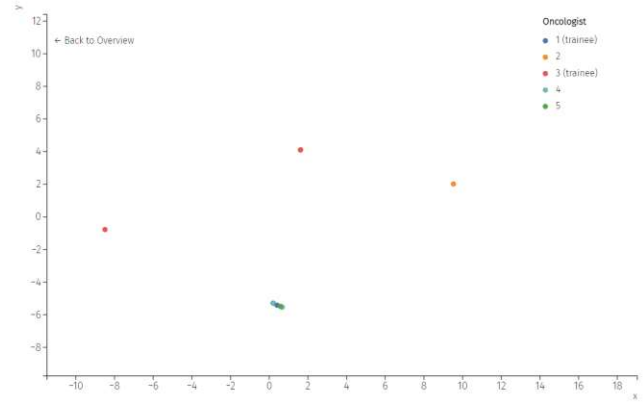
Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.



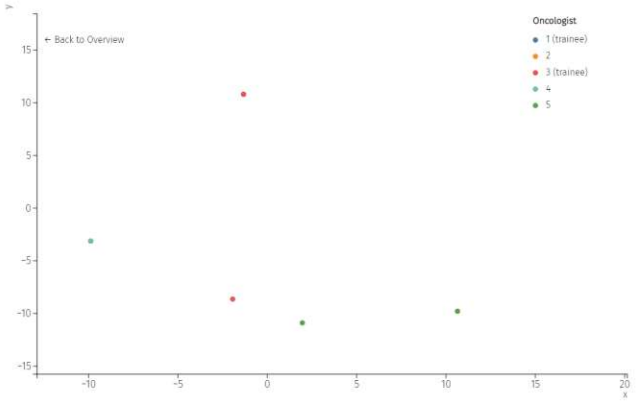
(a) `Interobs14` (without a selected probability threshold)



(b) `Interobs31` (without a selected probability threshold)

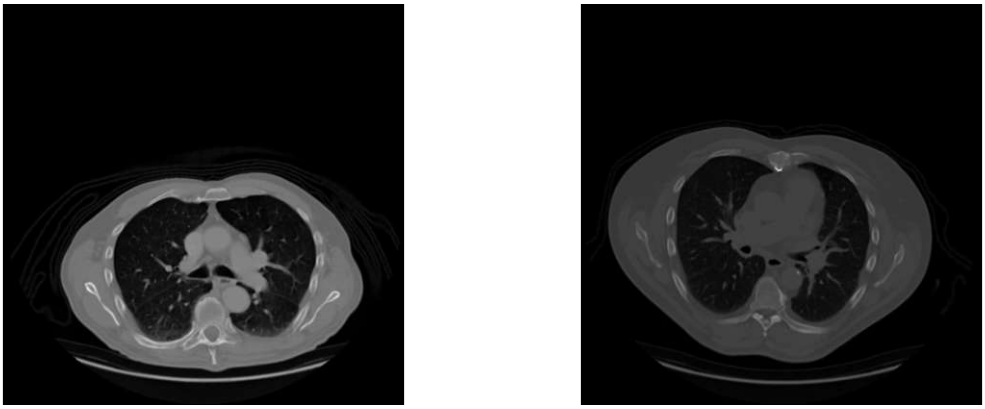


(c) `Interobs14` (60% segmentation probability)



(d) `Interobs31` (60% segmentation probability)

Figure 5.13: Comparison of segmentation-based radiomic features (after dimensionality reduction between `interobs14` (low-sensitivity patient) and `interobs31` (high-sensitivity patient) with respect to segmentation probability).



(a) `interobs14` (low-sensitivity patient) (b) `interobs31` (high-sensitivity patient)

Figure 5.14: Comparison of CT scans between a patients with different sensitivity regarding segmentation accuracy.

Aside from facilitating the analysis, the visualizations enabled a deep insight into radiomics data. By this, we refer to the visual interfaces fixating different subject matters, such as cluster analysis or the correlation of radiomics. For the cluster analysis, we provided information about the cluster distribution including size and validity of each cluster in a color-encoded dendrogram. We further determined the SHAP features of each cluster and visualized the feature value distributions for every patient in a selected cluster. Concerning the visualization of correlations within radiomics, we designed a scatterplot which is color-encoded based on clinical data. As mentioned previously, we preprocessed the radiomics data before visualizing these to ensure simplicity. The broad spectrum of view points in our radiomics analysis encourages a better understanding of data and underlying patterns. Furthermore, correlations between several aspects could be thoroughly inspected, which again helped us to uncover especially unexpected connections and outcomes.

Our approach helped us in confirming expected results. These encompass our anticipation of probabilistic segmentation heavily influencing the radiomics analysis as well as certain clinical data not necessarily correlating with clustering. Usage Scenarios 1, 2, 4, and 5 touch upon the topic of interobserver variability and segmentation accuracy significantly impacting the radiomics analysis. The findings presented in Usage Scenario 3 match our expectation of patient category “histology” not explicitly correlating with patient clusters in our case.

More importantly, our application enabled us to reveal unexpected patterns and outcomes. In Usage Scenario 1, we unveiled the numbers of clusters not increasing continuously with higher segmentation accuracy. We also detected similar findings for the cluster silhouette values with usage scenarios 2. Usage Scenario 4 showed us that cluster formation is mostly not coherent regarding different probability segmentation thresholds. In Usage Scenario 5, we revealed that some patient are more sensitive to the segmentation thresholds than

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

others, which is especially evident in the impact on the cluster analysis.

To sum up, *ProSeRa* supports the radiomics analysis by means of clear and insightful visualizations which represent radiomics data from various aspects. It considers clinical information in the analysis and further enables exploring the impact of different tumor segmentation probabilities on radiomics. Even though our tool enables us to confirm anticipated outcomes, we stress that our results should be interpreted in consideration to the limitations of our work. For example, one important aspect may be the current usage of a small data set. In the following section, we discuss the limitations of this thesis as well as the potential influence of these on our results in detail.

5.3 Limitations

As with the majority of research, the design of our visual analytics tool is subject to limitations. We identified the following shortcomings: (1) the small size of the input data set, (2) the omission of applying image filters prior to radiomics feature extraction, and (3) the lack of visualizations for representing tumor segmentations. We perceived these to have the highest potential impact on our findings and also provide interesting points for extensions in the future.

The small data set used for the radiomics analysis may have produced results which have strongly been influenced by the patient cohort and the group of oncologists. The 21 patients may not be representative of NSCLC patients, just the same as the segmentations by the five oncologists may not be representative of interobserver variability among radiation oncologist. This would make the results of our radiomics analysis being biased due to the subjectivity of oncologists as well as the chosen patient cohort (selection bias). However, our approach is flexible and additional or different data sets can be easily plugged in and analyzed with our dashboard.

The missing application of filters on medical images represents another limitation of our research. The usage of image filters prior to radiomics feature extraction may have enhanced the performance of calculation algorithms. This again could have resulted in obtaining outcomes which are more precise. This action would have been especially profitable for medical images which exhibit poor quality such as low contrast, low resolution or noise.

In contrast to the mentioned limitations which mostly concern the radiomics analysis, the lack of visualizations of tumor delineations on medical images relates to the analysis of probabilistic segmentation. Visual representations of all tumor segmentations conducted by each oncologist may have been beneficial for the investigation of interobserver variability and comprehension of deviations in tumor delineations. These visualizations could have enabled a better understanding of including specific (tumor) regions impacting the radiomics analysis. Such approaches have been previously investigated, for example by Jalalifar et al. [JSSSN22].

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

CHAPTER 6

Conclusion and Future Challenges

This thesis investigated to what extent the variability in tumor segmentations and the consideration of segmentation probability influence the outcome of radiomics analysis. We answer the research question formulated in Section 1.1 as follows:

> *"The variability in probabilistic tumor segmentations significantly influences the radiomics analysis especially in terms of clustering, which produces varying results among different segmentation probability thresholds."*

We come to this conclusion based on the results of our visual analytics tool *ProSeRa*, which consists of interactive visualizations representing radiomics data from various aspects. We designed and developed *ProSeRa* based on three aims. We fulfill the aim of visualizing and analyzing radiomics features with respect to clinical data (Aim A) by providing interactive visualizations which show radiomics data from various aspects, including cluster analysis and correlation. In doing so, we provide color encodings to convey information about the patient cohort as well as about cluster validity. In order to explore the effect of probabilistic segmentation on radiomics (Aim B), we used a data set containing multiple tumor delineations from different radiation oncologists. We preprocessed the extracted radiomics data resulting from each segmentation accuracy threshold in order to extract meaningful information prior to visualizing these. By providing interaction techniques such as *brushing and linking* and *focus-plus-context*, we address user interaction within our tool (Aim C). Based on the evaluation of our results using VDAR and QRI, we conclude that probabilistic tumor segmentations have a considerable impact on the radiomics analysis. Especially the cluster analysis was affected by the consideration of different segmentation accuracies as demonstrated in our usage scenarios.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Overall, our results show that tumor segmentations are a crucial factor for the radiomics analysis outcome. We have proven this by extracting radiomic features, reducing the dimensionality of these, performing cluster analysis, and computing the probabilistic tumor segmentations. We visualized our generated radiomics data on appropriate visualization types combined with the purposeful usage of colors. We thereby prevented information overflow on the visual interfaces in order to enable a clear analysis of radiomics information with respect to probability segmentation thresholds. We presented our results by means of five usage scenarios on remarkable findings. Our visual analytics tool greatly helped us in both confirming expected outcomes and unveiling unexpected insights into radiomics with regard to segmentation accuracy. For instance, the latter applies to cluster formation not being coherent regarding different probability segmentation thresholds.

Nonetheless, the results of our thesis should be regarded in light of its limitations. These mainly concern the impact on the radiomics analysis performance due to small sample size or omission of image filters prior to feature extraction. We also considered the missing visualizations for representing tumor delineations to be potentially disadvantageous for understanding the effect of including specific tumor regions.

Future work should specifically address the limitations of our current research to gain more insight on the correlation between probabilistic tumor segmentations and radiomics. This would include considering bigger data sets for the radiomics analysis. In order to further improve the performance of the analysis, the usage of suitable image filters prior to feature extraction should also be explored. Additionally, a more flexible way of conducting the cluster analysis should enable a more pliable exploration and analysis of the data set. For example, weighting factors should be taken into account for clustering when including, excluding, or weighting (clinical) features. In this case, guidance [CAA+20] and provenance [XAJK+15] could also be useful to support an easy navigation of the data and the entire analytical process. Finally, a more comprehensive probabilistic analysis of segmentations [PRW11, MCC+21] would provide additional insights, but this would require different dimensionality reduction and clustering approaches. For a better understanding of correlations between specific tumor regions and radiomics analysis, visualizations of tumor delineations should likewise be considered for future implementations. The consideration of these suggestions in future work could further extend the functionalities of *ProSeRa*, which already presents effective visualizations for investigating the effect of different tumor segmentations on the radiomics analysis.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

List of Figures

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

List of Tables

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

Acronyms

2D two-dimensional. 7, 13, 16, 21, 45

3D three-dimensional. 7, 13, 45, 67

BFT Belief Function Theory. 15

CNN convolutional neural networks. 17

CSV comma-separated values. 32, 45–47

CT Computed Tomography. 7, 15, 31, 32, 52, 72

DBSCAN density-based spatial clustering of applications with noise. 47

DICOM Digital Imaging and Communications in Medicine. 32, 45

EVEII Estimation of Imperfect Information. 15

FCM Fuzzy C-Means. 15, 16, 29

GLCM Gray-Level Coocurrence Matrix. 12, 45

GLDM Gray-Level Dependence Matrix. 65–67, 69, 70, 81

GLDZM Gray-Level Distance Zone Matrix. 12

GLRLM Gray-Level Run-Length Matrix. 12, 67, 69, 70

GLSZM Gray-Level Size Zone Matrix. 12, 45, 69

GTV gross tumor volume. 32

IBSI Image Biomarker Standardization Initiative. 8

KL Kullback-Leibler. 9

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

MDS Multi-dimensional Scaling. 9

MRI Magnetic Resonance Imaging. 7, 15

NGLDM Neighborhood Gray-Level Dependence Matrix. 12

NGTDM Neighborhood Gray-Tone Difference Matrix. 12

NRRD nearly raw raster data. 45, 49, 50

NSCLC non-small cell lung cancer. 31, 32, 56, 62, 66, 75

PCA Principal Component Analysis. 9, 46, 48

PCP parallel coordinate plots. 22–24, 28, 37, 79, 80

PDF probability density function. 49

PET Positron Emission Tomography. 7

QRI Qualitative Result Inspection. 3, 55, 77

ROI Region of Interest. 7, 8, 12, 13, 16, 17

SHAP SHapley Additive exPlanations. 34–36, 41, 44, 49, 51, 53, 65–67, 69, 70, 74, 80–83

SPLOM scatterplot matrix. 21, 79

t-SNE t-distributed Stochastic Neighbor Embedding. 9, 46, 48

TCGA The Cancer Genome Atlas. 9

TCIA The Cancer Imaging Archive. 3, 9, 31

UI user interface. 33–36, 80

UMAP Uniform Manifold Approximation and Projection. 46–48

VDAR Visual Data Analysis and Reasoning. 3, 55, 77

VOI Volume of Interest. 7, 8, 12, 13, 16

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Bibliography

[AK89] Moses Amadasun and Robert King. Textural Features Corresponding to Textural Properties. *IEEE Transactions on Systems, Man, and Cybernetics*, 19(5):1264–1274, 1989.

[BG97] Ingwer Borg and Patrick J. F. Groenen. *Modern Multidimensional Scaling*. Springer Series in Statistics. Springer, New York, USA, 1997.

[BKG+14] Yoganand Balagurunathan, Virendra Kumar, Yuhua Gu, Jongphil Kim, Hua Wang, Ying Liu, Dmitry B. Goldgof, Lawrence O. Hall, Rene Korn, Binsheng Zhao, Lawrence H. Schwartz, Satrajit Basu, Steven Eschrich, Robert A. Gatenby, and Robert J. Gillies. Test–Retest Reproducibility Analysis of Lung CT Image Features. *Journal of Digital Imaging*, 27(6):805–823, 2014.

[Bre01] Leo Breiman. Random Forests. *Machine Learning*, 45(1):5–32, 2001.

[CAA+20] Davide Ceneda, Natalia Andrienko, Gennady Andrienko, Theresia Gschwandtner, Silvia Miksch, Nikolaus Piccolotto, Tobias Schreck, Marc Streit, Josef Suschnigg, and Christian Tominski. Guide Me in Analysis: A Framework for Guidance Designers. *Computer Graphics Forum*, 39(6):269–288, 2020.

[CK19] Elisa Chotzoglou and Bernhard Kainz. Exploring the Relationship Between Segmentation Uncertainty, Segmentation Performance and Inter-observer Variability with Probabilistic Networks. In Luping Zhou, Nicholas Heller, Yiyu Shi, Yiming Xiao, Raphael Sznitman, Veronika Cheplygina, Diana Mateus, Emanuele Trucco, X. Sharon Hu, Danny Chen, Matthieu Chabanas, Hassan Rivaz, and Ingerid Reinertsen, editors, *Large-Scale Annotation of Biomedical Data and Expert Label Synthesis and Hardware Aware Learning for Medical Imaging and Computer Assisted Intervention*, pages 51–60, Cham, 10 2019. Springer International Publishing.

[CMS99] Stuart K. Card, Jock D. Mackinlay, and Ben Shneiderman. Readings in Information Visualization: Using Vision to Think (Interactive Technologies). *Morgan Kaufmann Publishers*, pages 295–305, 1999.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[CMT+19] Grzegorz Chlebus, Hans Meine, Smita Thoduka, Nasreddin Abolmaali, Bram van Ginneken, Horst Karl Hahn, and Andrea Schenk. Reducing inter-observer variability and interaction time of MR liver volumetry by combining automatic CNN-based liver segmentation and manual corrections. *PLOS ONE*, 14(5):1–14, 05 2019.

[Dan14] Priscille Dando. *Say It with Data: A Concise Guide to Making Your Case and Getting Results*. ALA Editions, Chicago, Illinois, USA, 2014.

[DS16] Deepa and Akansha Singh. Review of Brain Tumor Detection from MRI Images. In *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, pages 3997–4000, 2016.

[DSD+21] Kenny Davila, Srirangaraj Setlur, David Doermann, Bhargava Urala Kota, and Venu Govindaraju. Chart Mining: A Survey of Methods for Automated Chart Analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 43(11):3799–3819, 2021.

[EK80] Jenny Ellert and Louis Kreel. The Role of Computed Tomography in the Initial Staging and Subsequent Management of the Lymphomas. *Journal of Computer Assisted Tomography*, 4(3):368 – 391, 1980.

[EKSX96] Martin Ester, Hans-Peter Kriegel, Jörg Sander, and Xiaowei Xu. A Density-Based Algorithm for Discovering Clusters in Large Spatial Databases with Noise. In *Proceedings of the Second International Conference on Knowledge Discovery and Data Mining*, KDD'96, page 226–231. AAAI Press, 1996.

[FGL+16] Tim Furche, Georg Gottlob, Leonid Libkin, Giorgio Orsi, and Norman W. Paton. Data Wrangling for Big Data: Challenges and Opportunities. In *International Conference on Extending Database Technology*, 2016.

[Gal75] Mary M. Galloway. Texture Analysis Using Gray Level Run Lengths. *Computer Graphics and Image Processing*, 4(2):172–179, 1975.

[GDKB17] Ievgeniia Gutenko, Konstantin Dmitriev, Arie E. Kaufman, and Matthew A. Barish. AnaFe: Visual Analytics of Image-derived Temporal Features—Focusing on the Spleen. *IEEE Transactions on Visualization and Computer Graphics*, 23(1):171–180, 2017.

[Haj04] Materka Hajek. Texture Analysis Methodologies for Magnetic Resonance Imaging. *Dialogues in Clinical Neuroscience*, 6(2):243–250, 2004.

[Hal18] Maria Halkidi. Hierarchical Clustering. In Ling Liu and M. Tamer Özsu, editors, *Encyclopedia of Database Systems*, pages 1684–1689, New York, NY, USA, 2018. Springer New York.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[HE03] Greg Hamerly and Charles Elkan. Learning the k in k-means. In S. Thrun, L. Saul, and B. Schölkopf, editors, *Advances in Neural Information Processing Systems*, volume 16. MIT Press, 2003.

[HLD+17] Qiao Huang, Lin Lu, Laurent Dercle, Philip M.D. Lichtenstein, Yajun Li, Qian Yin, Min Zong, Lawrence Schwartz, and Binsheng Zhao. Interobserver variability in tumor contouring affects the use of radiomics to predict mutational status. *Journal of Medical Imaging*, 5(1):1–9, 2017.

[HR07] Jeffrey Heer and George Robertson. Animated Transitions in Statistical Data Graphics. *IEEE Transactions on Visualization and Computer Graphics*, 13(6):1240–1247, 2007.

[HS54] Brian Hopkins and J. G. Skellam. A New Method for determining the Type of Distribution of Plant Individuals. *Annals of Botany*, 18(2):213–227, 1954.

[HSD73] Robert M. Haralick, Kalaivani Shanmugam, and Its'Hak Dinstein. Textural Features for Image Classification. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-3(6):610–621, 1973.

[ID90] Alfred Inselberg and Bernard Dimsdale. Parallel Coordinates: A Tool for Visualizing Multi-dimensional Geometry. In *Proceedings of the First IEEE Conference on Visualization: Visualization '90*, pages 361–378, 1990.

[IIC+13] Tobias Isenberg, Petra Isenberg, Jian Chen, Michael Sedlmair, and Torsten Möller. A Systematic Review on the Practice of Evaluating Visualization. *IEEE Transactions on Visualization and Computer Graphics*, 19(12):2818–2827, 2013.

[Ins85] Alfred Inselberg. The Plane with Parallel Coordinates. *The Visual Computer*, 1(4), 1985.

[JME+18] Alain Jungo, Raphael Meier, Ekin Ermis, Marcela Blatti-Moreno, Evelyn Herrmann, Roland Wiest, and Mauricio Reyes. On the Effect of Inter-observer Variability for a Reliable Estimation of Uncertainty of Medical Image Segmentation. In Alejandro F. Frangi, Julia A. Schnabel, Christos Davatzikos, Carlos Alberola-López, and Gabor Fichtinger, editors, *Medical Image Computing and Computer Assisted Intervention – MICCAI 2018*, volume 11070, pages 682–690, Cham, 2018. Springer International Publishing.

[Jol02] Ian T. Jolliffe. *Principal Component Analysis*. Springer Series in Statistics. Springer, New York, NY, 2 edition, 2002.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[JSSSN22] Seyed Ali Jalalifar, Hany Soliman, Arjun Sahgal, and Ali Sadeghi-Naini. Impact of tumour segmentation accuracy on efficacy of quantitative mri biomarkers of radiotherapy outcome in brain metastasis. *Cancers*, 14(20), 2022.

[Kei02] Daniel A. Keim. Information Visualization and Visual Data Mining. *IEEE Transactions on Visualization and Computer Graphics*, 8(1):1–8, 2002.

[Kro08] Anders Krogh. What are artificial neural networks? *Nature Biotechnology*, 26(2):195–197, 2008.

[Kru64a] Joseph B. Kruskal. Multidimensional scaling by optimizing goodness of fit to a nonmetric hypothesis. *Psychometrika*, 29:1–27, 1964.

[Kru64b] Joseph B. Kruskal. Nonmetric multidimensional scaling: A numerical method. *Psychometrika*, 29:115–129, 1964.

[Kul68] Solomon Kullback. *Information Theory and Statistics*. Dover Publications Inc., Mineola, New York, USA, 1968.

[LBI+12] Heidi Lam, Enrico Bertini, Petra Isenberg, Catherine Plaisant, and Sheelagh Carpendale. Empirical Studies in Information Visualization: Seven Scenarios. *IEEE Transactions on Visualization and Computer Graphics*, 18(9):1520–1536, 2012.

[LCK+14] Timothy J. Larkin, Holly C. Canuto, Mikko I. Kettunen, Thomas C. Booth, De-En Hu, Anant S. Krishnan, Sarah E. Bohndiek, André A. Neves, Charles McLachlan, Michael P. Hobson, and Kevin M. Brindle. Analysis of image heterogeneity using 2D Minkowski functionals detects tumor responses to treatment. *Magnetic Resonance in Medicine*, 71(1):402–410, 2014.

[LGM+14] Benoit Lelandais, Isabelle Gardin, Laurent Mouchard, Pierre Vera, and Su Ruan. Dealing with uncertainty and imprecision in image segmentation using belief function theory. *International Journal of Approximate Reasoning*, 55(1, Part 3):376–387, 2014. Theory and applications of belief functions – Belief 2012.

[LL17] Scott M. Lundberg and Su-In Lee. A Unified Approach to Interpreting Model Predictions. In *Proceedings of the 31st International Conference on Neural Information Processing Systems*, NIPS'17, page 4768–4777, Red Hook, NY, USA, 2017. Curran Associates Inc.

[LPL+22] Xiao Luo, Sirong Piao, Haiqing Li, Yuxin Li, Wei Xia, Yifang Bao, Xueling Liu, Daoying Geng, Hao Wu, and Liqin Yang. Multi-lesion radiomics model for discrimination of relapsing-remitting multiple

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.

sclerosis and neuropsychiatric systemic lupus erythematosus. *European Radiology*, 32(8):5700–5710, 2022.

[LRVL+12] Philippe Lambin, Emmanuel Rios-Velazquez, Ralph Leijenaar, Sara Carvalho, Ruud G.P.M. van Stiphout, Patrick Granton, Catharina M.L. Zegers, Robert Gillies, Ronald Boellard, André Dekker, and Hugo J.W.L.g Aerts. Radiomics: Extracting more information from medical images using advanced feature analysis. *European Journal of Cancer*, 48(4):441–446, 2012.

[LvTdJ+17] Ruben T. H. M. Larue, Janna E. van Timmeren, Evelyn E. C. de Jong, Giacomo Feliciani, Ralph T. H. Leijenaar, Wendy M. J. Schreurs, Meindert N. Sosef, Frank H. P. J. Raat, Frans H. R. van der Zande, Marco Das, Wouter van Elmpt, and Philippe Lambin. Influence of gray level discretization on radiomic feature stability for different CT scanners, tube currents and slice thicknesses: a comprehensive phantom study. *Acta Oncologica*, 56(11):1544–1553, 2017.

[MCC+21] Ji Ma, Jinjin Chen, Liye Chen, Jiazhou Chen, Xujia Qin, and Mingyu Bao. Probabilistic Slider: A Tool for Visualizing Fuzzy Segmentation Uncertainties. *IEEE Access*, 9:28707–28715, 2021.

[MEH+22] Eric Mörth, Tanja Eichner, Ingfrid Haldorsen, Stefan Bruckner, and Noeska N. Smit. ICEVis: Interactive Clustering Exploration for Tumor Sub-Region Analysis in Multiparametric Cancer Imaging. In *Proceedings of the 15th International Symposium on Visual Information Communication and Interaction*, VINCI '22, New York, NY, USA, 2022. Association for Computing Machinery.

[MHM18] Leland McInnes, John Healy, and James Melville. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction, 2018.

[MML+20] Marius E. Mayerhoefer, Andrzej Materka, Georg Langs, Ida Häggström, Piotr Szczypiński, Peter Gibbs, and Gary Cook. Introduction to Radiomics. *Journal of Nuclear Medicine*, 61(4):488–495, 2020.

[Mun09] Tamara Munzner. A Nested Model for Visualization Design and Validation. *IEEE Transactions on Visualization and Computer Graphics*, 15(6):921–928, 2009.

[MWLH+20] Eric Mörth, Kari Wagner-Larsen, Erlend Hodneland, Camilla Krakstad, Ingfrid S. Haldorsen, Stefan Bruckner, and Noeska N. Smit. RadEx: Integrated Visual Exploration of Multiparametric Studies for Radiomic Tumor Profiling. *Computer Graphics Forum*, 2020.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[NMI10] Samina Naz, Hammad Majeed, and Humayun Irshad. Image Segmentation using Fuzzy Clustering: A Survey. In *Emerging Technologies (ICET), 2010 6th International Conference on*, pages 181–186, 2010.

[PHPQACPQ20] Ronald M. Parra-Hernández, Jorge I. Posada-Quintero, Orlando Acevedo-Charry, and Hugo F. Posada-Quintero. Uniform Manifold Approximation and Projection for Clustering Taxa through Vocalizations in a Neotropical Passerine (Rough-Legged Tyrannulet, Phyllomyias burmeisteri). *Animals*, 10(8), 2020.

[PLG+15] Chintan Parmar, Ralph T. H. Leijenaar, Patrick Grossmann, Emmanuel Rios Velazquez, Johan Bussink, Derek Rietveld, Michelle M. Rietbergen, Benjamin Haibe-Kains, Philippe Lambin, and Hugo J.W.L. Aerts. Radiomic feature clusters and Prognostic Signatures specific for Lung and Head & Neck cancer. *Scientific Reports*, 5(1), 2015.

[PRW11] Tobias Pfaffelmoser, Matthias Reitinger, and Rüdiger Westermann. Visualizing the Positional and Geometrical Variability of Isosurfaces in Uncertain Scalar Fields. *Computer Graphics Forum*, 30(3):951–960, 2011.

[PS17] Madhuri Pawar and Deepali Sale. MRI and CT Image Denoising using Gaussian Filter, Wavelet Transform and Curvelet Transform. *International Journal of Engineering Science and Computing*, 7(5), 2017.

[PT01] Regina Pohle and Klaus D. Toennies. Segmentation of Medical Images Using Adaptive Region Growing. In Milan Sonka and Kenneth M. Hanson, editors, *Medical Imaging 2001: Image Processing*, volume 4322, pages 1337 – 1346. International Society for Optics and Photonics, SPIE, 2001.

[PVG+11] Fabian Pedregosa, Gaël Varoquaux, Alexandre Gramfort, Vincent Michel, Bertrand Thirion, Olivier Grisel, Mathieu Blondel, Peter Prettenhofer, Ron Weiss, Vincent Dubourg, Jake Vanderplas, Alexandre Passos, David Cournapeau, Matthieu Brucher, Matthieu Perrot, and Édouard Duchesnay. Scikit-learn: Machine Learning in Python. *Journal of Machine Learning Research*, 12:2825–2830, 2011.

[RBR+18] Stefania Rizzo, Francesca Botta, Sara Raimondi, Daniela Origgi, Cristiana Fanciullo, Alessio Morganti, and Massimo Bellomi. Radiomics: the facts and the challenges of image analysis. *European Radiology Experimental*, 2(1):36, 12 2018.

[RFB15] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-Net: Convolutional Networks for Biomedical Image Segmentation. In Nassir

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

Navab, Joachim Hornegger, William M. Wells, and Alejandro F. Frangi, editors, *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, pages 234–241, Cham, 2015. Springer International Publishing.

[RG21] Rahul and Bhawna Goyal. Gaussian Filtering based Image Integration for Improved Disease Diagnosis and Treatment Planning. In *2021 9th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions) (ICRITO)*, pages 1–6, 2021.

[Rou87] Peter J. Rousseeuw. Silhouettes: A graphical aid to the interpretation and validation of cluster analysis. *Journal of Computational and Applied Mathematics*, 20:53–65, 1987.

[RPB+16] Stefania Rizzo, Francesco Petrella, Valentina Buscarino, Federica De Maria, Sara Raimondi, Massimo Barberis, Caterina Fumagalli, Gianluca Spitaleri, Cristiano Rampinelli, Filippo De Marinis, Lorenzo Spaggiari, and Massimo Bellomi. CT Radiogenomic Characterization of EGFR, K-RAS, and ALK Mutations in Non-Small Cell Lung Cancer. *European Radiology*, 26(1):32–42, 2016.

[RTB96] Bernice E. Rogowitz, Lloyd A. Treinish, and Steve Bryson. How Not to Lie with Visualization. *Computers in Physics*, 10(3):268–273, 1996.

[RvdHD+15] Renata Georgia Raidou, U.A. van der Heide, C. V. Dinh, G. Ghobadi, J. F. Kallehauge, M. Breeuwer, and Anna Vilanova. Visual Analytics for the Exploration of Tumor Tissue Characterization. *Eurographics Conference on Visualization (EuroVis) 2015*, 34(3):11–20, 2015.

[SGH+16] Ashirbani Saha, Lars J. Grimm, Michael Harowicz, Sujata V. Ghate, Connie Kim, Ruth Walsh, and Maciej A. Mazurowski. Interobserver variability in identification of breast tumors in MRI and its implications for prognostic biomarkers and radiogenomics. *Medical Physics*, 43(8):4558–4564, 2016.

[SHB10] Astrid Schneider, Gerhard Hommel, and Maria Blettner. Linear Regression Analysis. *Deutsches Ärzteblatt International*, 107(44):776–782, 2010.

[SMH10] Ahmed Saad, Torsten Möller, and Ghassan Hamarneh. ProbExplorer: Uncertainty-guided Exploration and Editing of Probabilistic Medical Image Segmentation. *Computer Graphics Forum*, 29(3):1113–1122, 2010.

[Spe07] Robert Spence. *Information Visualization: Design for Interaction*. Pearson education. Pearson/Prentice Hall, 2007.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[SSM11] Samuel Silva, Beatriz Sousa Santos, and Joaquim Madeira. Using color in visualization: A survey. *Computers & Graphics*, 35(2):320–333, 2011.

[SSM+16] Matthew Sadiku, Adebowale Shadare, Sarhan Musa, Cajetan Akujuobi, and Roy Perry. Data Visualization. *International Journal of Engineering Research and Advanced Technology (IJERAT)*, 12:2454–6135, 12 2016.

[SSP+22] Isaac Shiri, Yazdan Salimi, Masoumeh Pakbin, Ghasem Hajianfar, Atlas Haddadi Avval, Amirhossein Sanaat, Shayan Mostafaei, Azadeh Akhavanallaf, Abdollah Saberi, Zahra Mansouri, Dariush Askari, Mohammadreza Ghasemian, Ehsan Sharifipour, Saleh Sandoughdaran, Ahmad Sohrabi, Elham Sadati, Somayeh Livani, Pooya Iranpour, Shahriar Kolahi, Maziar Khateri, Salar Bijari, Mohammad Reza Atashzar, Sajad P. Shayesteh, Bardia Khosravi, Mohammad Reza Babaei, Elnaz Jenabi, Mohammad Hasanian, Alireza Shahhamzeh, Seyaed Yaser Foroghi Ghomi, Abolfazl Mozafari, Arash Teimouri, Fatemeh Movaseghi, Azin Ahmari, Neda Goharpey, Rama Bozorgmehr, Hesamaddin Shirzad-Aski, Roozbeh Mortazavi, Jalal Karimi, Nazanin Mortazavi, Sima Besharat, Mandana Afsharpad, Hamid Abdollahi, Parham Geramifar, Amir Reza Radmard, Hossein Arabi, Kiara Rezaei-Kalantari, Mehrdad Oveisi, Arman Rahmim, and Habib Zaidi. COVID-19 prognostic modeling using CT radiomic features and machine learning algorithms: Analysis of a multi-institutional dataset of 14,339 patients. *Computers in Biology and Medicine*, 145:105467, 2022.

[SV16] N. Senthilkumaran and S. Vaithegi. Image Segmentation By Using Thresholding Techniques For Medical Images. *Computer Science & Engineering: An International Journal*, 6:1–13, 02 2016.

[SvCT+20] Martijn P.A. Starmans, Sebastian R. van der Voort, Jose M. Castillo Tovar, Jifke F. Veenland, Stefan Klein, and Wiro J. Niessen. Chapter 18 - Radiomics: Data mining using quantitative medical image features. In S. Kevin Zhou, Daniel Rueckert, and Gabor Fichtinger, editors, *Handbook of Medical Image Computing and Computer Assisted Intervention*, The Elsevier and MICCAI Society Book Series, pages 429–456. Academic Press, 2020.

[SW83] Chengjun Sun and William G. Wee. Neighboring gray level dependence matrix for texture classification. *Computer Vision, Graphics, and Image Processing*, 23(3):341–352, 1983.

[SWRT14] Michaela Spitzer, Jan Wildenhain, Juri Rappsilber, and Mike Tyers. BoxPlotR: A web tool for generation of box plots. *Nature Methods*, 11(2):121–122, 2014.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[TAM14] Guillaume Thibault, Jesús Angulo, and Fernand Meyer. Advanced Statistical Matrices for Texture Characterization: Application to Cell Classification. *IEEE Transactions on Biomedical Engineering*, 61(3):630–637, 2014.

[TBN+15] William D. Travis, Elisabeth Brambilla, Andrew G. Nicholson, Yasushi Yatabe, John H.M. Austin, Mary Beth Beasley, Lucian. R. Chirieac, Sanja Dacic, Edwina Duhig, Douglas B. Flieder, Kim Geisinger, Fred R. Hirsch, Yuichi Ishikawa, Keith M. Kerr, Masayuki Noguchi, Giuseppe Pelosi, Charles A. Powell, Ming Sound Tsao, and Ignacio Wistuba. The 2015 World Health Organization Classification of Lung Tumors: Impact of Genetic, Clinical and Radiologic Advances Since the 2004 Classification. *Journal of Thoracic Oncology*, 10(9):1243–1260, 2015.

[TMB+18] Rajat Thawani, Michael McLane, Niha Beig, Soumya Ghose, Prateek Prasanna, Vamsidhar Velcheti, and Anant Madabhushi. Radiomics and radiogenomics in lung cancer: A review for the clinician. *Lung Cancer*, 115:34–41, 2018.

[TSK05] Pang-Ning Tan, Michael Steinbach, and Vipin Kumar. *Introduction to Data Mining*. Addison Wesley, 2005.

[vdMH08] Laurens van der Maaten and Geoffrey E. Hinton. Visualizing Data using t-SNE. *Journal of Machine Learning Research*, 9(86):2579–2605, 2008.

[vGFP+17a] Joost J.M. van Griethuysen, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina G.H. Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo J.W.L. Aerts. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, 77(21):e104–e107, 2017.

[vGFP+17b] Joost J.M. van Griethuysen, Andriy Fedorov, Chintan Parmar, Ahmed Hosny, Nicole Aucoin, Vivek Narayan, Regina G.H. Beets-Tan, Jean-Christophe Fillion-Robin, Steve Pieper, and Hugo J.W.L. Aerts. Computational Radiomics System to Decode the Radiographic Phenotype. *Cancer Research*, 77(21):104–107, 10 2017.

[vTCTL+20] Janita E. van Timmeren, Davide Cester, Stephanie Tanadini-Lang, Hatem Alkadhi, and Bettina Baessler. Radiomics in medical imaging — "how-to" guide and critical reflection. *Insights into Imaging*, 11(1):91, 2020.

[WAKD19] Leonard Wee, Hugo J.W.L. Aerts, Petros Kalendralis, and Andre Dekker. Data from NSCLC-Radiomics-Interobserver1 [Data set], 2019. The Cancer Imaging Archive. `https://doi.org/10.7937/tcia.2019.cwvlpd26`.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[WBW+21] Jeffrey Wong, Michael Baine, Sarah Wisnoskie, Nathan Bennion, Dechun Zheng, Lei Yu, Vipin Dalal, Michael A. Hollingsworth, Chi Lin, and Dandan Zheng. Effects of interobserver and interdisciplinary segmentation variabilities on CT-based radiomics for pancreatic cancer. *Scientific Reports*, 11(1), 2021.

[WD17] Ryan Wilson and Anand Devaraj. Radiomics of pulmonary nodules and lung cancer. *Translational Lung Cancer Research*, 6(1):86–91, 2017.

[WF09] Leland Wilkinson and Michael Friendly. The History of the Cluster Heat Map. *The American Statistician*, 63(2):179–184, 2009.

[WHL+20] Bianca Williams, Caroline Halloin, Wiebke Löbel, Ferdous Finklea, Elizabeth Lipke, Robert Zweigerdt, and Selen Cremaschi. Data-Driven Model Development for Cardiomyocyte Production Experimental Failure Prediction. In Sauro Pierucci, Flavio Manenti, Giulia Luisa Bozzano, and Davide Manca, editors, *30th European Symposium on Computer Aided Process Engineering*, volume 48 of *Computer Aided Chemical Engineering*, pages 1639–1644. Elsevier, 2020.

[WSJ+13] Takeyuki Watadani, Fumikazu Sakai, Takeshi Johkoh, Satoshi Noma, Masanori Akira, Kiminori Fujimoto, Alexander A. Bankier, Kyung Soo Lee, Nestor L. Müller, Jae-Woo Song, Jai-Soung Park, David A. Lynch, David M. Hansell, Martine Remy-Jardin, Tomás Franquet, and Yukihiko Sugiyama. Interobserver Variability in the CT Assessment of Honeycombing in the Lungs. *Radiology*, 266(3):936–944, 2013.

[WZZ+19] Jingjun Wu, Qinhe Zhang, Ying Zhao, Yijun Liu, Anliang Chen, Xin Li, Tingfan Wu, Jianying Li, Yan Guo, and Ailian Liu. Radiomics Analysis of Iodine-Based Material Decomposition Images With Dual-Energy Computed Tomography Imaging for Preoperatively Predicting Microsatellite Instability Status in Colorectal Cancer. *Frontiers in Oncology*, 9:1250, 11 2019.

[XAJK+15] Kai Xu, Simon Attfield, T.J. Jankun-Kelly, Ashley Wheat, Phong H. Nguyen, and Nallini Selvaraj. Analytic Provenance for Sensemaking: A Research Agenda. *IEEE Computer Graphics and Applications*, 35(3):56–64, 2015.

[YA16] Stephen S. F. Yip and Hugo J. W. L. Aerts. Applications and limitations of radiomics. *Physics in Medicine and Biology*, 61(13):150–166, 2016.

[YJY+17] Lina Yu, Hengle Jiang, Hongfeng Yu, Chi Zhang, Josiah Mcallister, and Dandan Zheng. iVAR: Interactive visual analytics of radiomics features from large-scale medical images. In *2017 IEEE International Conference on Big Data (Big Data)*, pages 3916–3923. IEEE, 2017.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.
TU Bibliothek Your knowledge hub WIEN

[ZVA+20] Alex Zwanenburg, Martin Vallières, Mahmoud A. Abdalah, Hugo J. W. L. Aerts, Vincent Andrearczyk, Aditya Apte, Saeed Ashrafinia, Spyridon Bakas, Roelof J. Beukinga, Ronald Boellaard, Marta Bogowicz, Luca Boldrini, Irène Buvat, Gary J. R. Cook, Christos Davatzikos, Adrien Depeursinge, Marie-Charlotte Desseroit, Nicola Dinapoli, Cuong Viet Dinh, Sebastian Echegaray, Issam El Naqa, Andriy Y. Fedorov, Roberto Gatta, Robert J. Gillies, Vicky Goh, Michael Götz, Matthias Guckenberger, Sung Min Ha, Mathieu Hatt, Fabian Isensee, Philippe Lambin, Stefan Leger, Ralph T.H. Leijenaar, Jacopo Lenkowicz, Fiona Lippert, Are Losnegård, Klaus H. Maier-Hein, Olivier Morin, Henning Müller, Sandy Napel, Christophe Nioche, Fanny Orlhac, Sarthak Pati, Elisabeth A.G. Pfaehler, Arman Rahmim, Arvind U.K. Rao, Jonas Scherer, Muhammad Musib Siddique, Nanna M. Sijtsema, Jairo Socarras Fernandez, Emiliano Spezi, Roel J.H.M. Steenbakkers, Stephanie Tanadini-Lang, Daniela Thorwarth, Esther G.C. Troost, Taman Upadhaya, Vincenzo Valentini, Lisanne V. van Dijk, Joost van Griethuysen, Floris H.P. van Velden, Philip Whybra, Christian Richter, and Steffen Löck. The Image Biomarker Standardization Initiative: Standardized Quantitative Radiomics for High-Throughput Image-based Phenotyping. *Radiology*, 295(2):328–338, 2020.

Die approbierte gedruckte Originalversion dieser Diplomarbeit ist an der TU Wien Bibliothek verfügbar
The approved original version of this thesis is available in print at TU Wien Bibliothek.